

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 18, 2003, 23:13:13 ; Search time 6955 Seconds

(without alignments)
11323.105 Million cell updates/sec

Title: US-09-920-705-2

Perfect score: 2706
Sequence: 1 atggcgtatctatgaagcc.....cgctgacagggccttga 2706Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Genembl.*
1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pt.*
10: gb_ro.*
11: gb_sy.*
12: gb_un.*
13: gb_vl.*
14: gb_vl.*
15: em_da.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*
29: em_vl.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pln.*
35: em_htg_rod.*
36: em_htg_mam.*
37: em_htg_vrt.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2706	100.0	2706	AX420420	AX420420 Sequence
2	2701.2	99.8	3498	AY094463	AY094463 Arabidops
3	1896.8	70.1	4593	AX420419	AX420419 Sequence
C 4	1896.8	70.1	102057	AC002335	AC002335 Arabidops
C 5	797.8	29.5	801	AX420422	AX420422 Sequence
6	271.4	10.0	349	AX420423	AX420423 Sequence
7	271.4	10.0	3715	AX420424	AX420424 Sequence
C 8	148	5.5	15345	AC108755	AC108755 Oryza sat
9	108.8	4.0	137519	ATM1P17	AL049730 Arabidops
C 10	108.8	4.0	196107	ATCHRV34	AL161534 Arabidops
11	87	3.2	132092	AP004011	AP004011 Oryza sat
C 12	47.6	1.8	7218	166494	166494 Sequence 14
13	45.4	1.7	161477	AC103481	AC103481 Rattus no
14	45.2	1.7	1141	AX083744	AX083744 Sequence
15	44	1.6	152079	AC094195	AC094195 Rattus no
16	42.8	1.6	256	AX261730	AX261730 Sequence
17	42.6	1.6	180418	AC106674	AC106674 Rattus no
18	42.2	1.6	1410	M0SCARGF	L36653 Mus musculu
19	42.2	1.6	1595	AX305833	AX305833 Sequence
20	42.2	1.6	1595	M0SCARG	D90151 Mus musculu
21	42.2	1.6	1893	AK001314	AK001314 Homo sapi
22	42.2	1.6	1961	BC002343	BC002343 Homo sapi
23	42.2	1.6	1966	BC006494	BC006494 Homo sapi
24	42.2	1.6	1966	BC006516	BC006516 Homo sapi
25	42.2	1.6	2312	AK000250	AK000250 Homo sapi
26	42.2	1.6	2542	AK000221	AK000221 Homo sapi
27	42.2	1.6	2560	AK095897	AK095897 Homo sapi
28	42.2	1.6	2637	AK091742	AK091742 Homo sapi
C 29	42.2	1.6	125020	AF429315	AF429315 Homo sapi
30	42	1.6	1978	RATRA301	D9708 Rattus norv
31	41.8	1.5	191603	AL713974	AL713974 Mouse DNA
32	41.6	1.5	579	M13757	M13757 Chinese ham
33	41.6	1.5	2142	CIUC23A	AX401891 Sequence
34	41.6	1.5	2229	CRUC23	M15823 Chinese ham
35	41.4	1.5	2052	AX305666	AX305666 Sequence
36	41.4	1.5	2052	MMS1G41	X80232 M.musculus
37	41.4	1.5	318930	AC073495	AC073495 Mus muscu
C 38	41.2	1.5	165720	AC025524	AC025524 Homo sapi
39	41	1.5	3205	AK092332	AK092332 Homo sapi
40	41	1.5	11145	HS1340191	AL096858 Novel hum
41	41	1.5	12227	AP356524	AP356524 Homo sapi
C 42	41	1.5	143501	AC104583	AC104583 Homo sapi
C 43	41	1.5	146836	AC019317	AC019317 Homo sapi
C 44	40.8	1.5	1141	AX083744	AX083744 Sequence
45	40.6	1.5	158174	AC123065	AC123065 Mus muscu

ALIGNMENTS

RESULT 1
AX420420
LOCUS AX420420
DEFINITION Sequence 2 from Patent WO0212518.
ACCESSION AX420420
VERSION AX420420.1 GI:21524576
KEYWORDS
SOURCE
ORGANISM
thale cress.
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

Pred. No. is the number of results predicted by chance to have a

REFERENCE
AUTHORS
TITLE
Amasino,R.M., Schomburg,F.M., Michaels,S.D. and Patton,D.
Floral induction gene

Dp	1561	CGGGGGGGGTGTGGCAAAATATGATGATGTGTCAACATTATTTCTTGTTGGCCCTCCATTCAGAT	1620
Qy	1621	TTCTTTAACTGATGTACTTCCAAAGTAGACCCGGTCAGAAGCGCTATATATGGTTGTTCTCAAG	1680
Dp	1621	TTCTTTAACTGATGTACTTCCAAAGTAGACCCGGTCAGAAGCGCTATATATGGTTGTTCTCAAG	1680
Qy	1661	TTTACCCCCCGCAGCCGTTCCCTGTTACAGCATCATACAGACAAGAAATCTCAATTCCT	1740
Dp	1661	TTTACCCCCCGCAGCCGTTCCCTGTTACAGCATCATACAGACAAGAAATCTCAATTCCT	1740
Qy	1741	CTGCATTTATATGATATCAAGCGCGGGAAATTCACCTCGCAATGCTATGTACAGTTTATATCCT	1800
Dp	1741	CTGCATTTATATGATATCAAGCGCGGGAAATTCACCTCGCAATGCTATGTACAGTTTATATCCT	1800
Qy	1801	CCTAGGGAATAATTACATTATGGGGGTGCACACAGAACATTTTGAACGCTGTTCAAAACCATCT	1860
Dp	1801	CCTAGGGAATAATTACATTATGGGGGTGCACACAGAACATTTTGAACGCTGTTCAAAACCATCT	1860
Qy	1861	GTTAGCGAGCCTCTCAGATATCCTTAATATATGACAGCGCCTCAAGCTGGGGTTAGTTTAAT	1920
Dp	1861	GTTAGCGAGCCTCTCAGATATCCTTAATATATGACAGCGCCTCAAGCTGGGGTTAGTTTAAT	1920
Qy	1921	CCGGAGCTTTTATAGCACACTCTGGCAATATTTCTCCCTGCAATTTCTCAACCGCTGGCCCT	1980
Dp	1921	CCGGAGCTTTTATAGCACACTCTGGCAATATTTCTCCCTGCAATTTCTCAACCGCTGGCCCT	1980
Qy	1961	GAGAGTCACCAACCTATGTGCAGGACCTTCAACAGTTGTTTTCACAGCACATCAAGTCCAAAT	2040
Dp	1961	GAGAGTCACCAACCTATGTGCAGGACCTTCAACAGTTGTTTTCACAGCACATCAAGTCCAAAT	2040
Qy	2041	GGACTGTATCAATTGGAGAAAGCACCGCTGTACAGCTTGGAAAGAGTCCCAAAACAGTTTCAT	2100
Dp	2041	GGACTGTATCAATTGGAGAAAGCACCGCTGTACAGCTTGGAAAGAGTCCCAAAACAGTTTCAT	2100
Qy	2101	GATCGTCGTAATTCAGTCACTTCCAAACAATAGCGAATATGATACGCTCCAGCGTGGGCACTA	2160
Dp	2101	GATCGTCGTAATTCAGTCACTTCCAAACAATAGCGAATATGATACGCTCCAGCGTGGGCACTA	2160
Qy	2161	CCTCCTCCTCTCTTGCGGCTTTTACCCCTCCAGCTTCAAAACAACCCCAACTACATAGTGGATG	2220
Dp	2161	CCTCCTCCTCTCTTGCGGCTTTTACCCCTCCAGCTTCAAAACAACCCCAACTACATAGTGGATG	2220
Qy	2221	GTCATATGGCAATATGCAATATCCACAAGCCAAATCTGTAAACATGAGCTGTCCGCTTA	2280
Dp	2221	GTCATATGGCAATATGCAATATCCACAAGCCAAATCTGTAAACATGAGCTGTCCGCTTA	2280
Qy	2281	CCAAATATGGCCTCAATTAATTAATTTTCATGTACACTCAGGGTGTGCTAAATCAATCCGTTT	2340
Dp	2281	CCAAATATGGCCTCAATTAATTAATTTTCATGTACACTCAGGGTGTGCTAAATCAATCCGTTT	2340
Qy	2341	TCTCAGCCCATATGTATCCAAATATGCAATATGCAATATGCAATATGCAATATGCAATATGCAAT	2400
Dp	2341	TCTCAGCCCATATGTATCCAAATATGCAATATGCAATATGCAATATGCAATATGCAATATGCAAT	2400
Qy	2401	CCAATTTCCAAATTTATCAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG	2460
Dp	2401	CCAATTTCCAAATTTATCAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG	2460
Qy	2461	TTTAAACCCCTTCCCAATTTTCAAGCGGCAATGCAACACACACAGAGATTAAGGCAAAATTTAGAG	2520
Dp	2461	TTTAAACCCCTTCCCAATTTTCAAGCGGCAATGCAACACACACAGAGATTAAGGCAAAATTTAGAG	2520
Qy	2521	CCACAAAAACCAAGCACTACGATTTGCAGCTTATGATCTCTGGGGATGGTCAAGGATACACA	2580
Dp	2521	CCACAAAAACCAAGCACTACGATTTGCAGCTTATGATCTCTGGGGATGGTCAAGGATACACA	2580
Qy	2561	GATGGGGAGTGTGATTAAGAAATCAAGATATACAGTCAACACTACATCAAAATTTGCAGCAAACTT	2640
Dp	2561	GATGGGGAGTGTGATTAAGAAATCAAGATATACAGTCAACACTACATCAAAATTTGCAGCAAACTT	2640
Qy	2641	CTTCTCCAGATATACAGCAAAACACACAGCAACAGCTTCAAGTACTCCGGCTGGACAGGG	2700
Dp	2641	CTTCTCCAGATATACAGCAAAACACACAGCAACAGCTTCAAGTACTCCGGCTGGACAGGG	2700

OY	2701	CCTTGA	2706	
Db	2701	CCTTGA	2706	
RESULT 2				
AY094463				
LOCUS	AY094463	3498 bp	mRNA	linear
DEFINITION	Arabidopsis thaliana At2g43410/T1024.15	mRNA, complete cds.		
ACCESSION	AY094463			
VERSION	AY094463.1	GI:20453184		
KEYWORDS	FLI CDNA.			
SOURCE	Arabidopsis thaliana.			
ORGANISM	Arabidopsis thaliana			
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.			
AUTHORS	1 (bases 1 to 3498)			
REFERENCE	Shinn, P., Chen, H., Cheuk, R., Kim, C.J., Meyers, M.C., Banh, J., Bowser, L., Carninci, P., Chang, E., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Wu, H.C., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.			
TITLE	Arabidopsis cDNA clones			
JOURNAL	Arabidopsis cDNA clones			
REFERENCE	2 (bases 1 to 3498)			
AUTHORS	Shinn, P., Chen, H., Cheuk, R., Kim, C.J., Meyers, M.C., Banh, J., Bowser, L., Carninci, P., Chang, E., Dale, J.M., Goldsmith, A.D., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Wu, H.C., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.			
TITLE	Direct Submission			
JOURNAL	Submitted (08-APR-2002) Salk Institute Genomic Analysis Laboratory (Stigall), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA			
COMMENT	RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL CDNA : 'RIKEN Arabidopsis Full-length cDNA') : Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.			
FEATURES	The Salk, Stanford, PGSC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Shinn, P., Chen, H., Cheuk, R., Kim, C.J., Meyers, M.C., Banh, J., Bowser, L., Chang, E., Dale, J.M., Goldsmith, A.D., Jones, T., Karlin-Neumann, G., Lam, B., Lee, J.M., Lin, J., Miranda, M., Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Southwick, A., Tang, C.C., Toriumi, M., Wu, H.C., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Davis, R.W., Theologis, A. and Ecker, J.R.			
CDS	Shinn, P. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk) contributed equally to this work as PIs.			
	Location/Qualifiers			
	1..3498			
	/organism="Arabidopsis thaliana"			
	/db_xref="taxon:3702"			
	/chromosome="2"			
	/clone="RAFL09-47-p10(R21966)"			
	/note="ecotype: Columbia"			
	255..2960			
	/note="unknown protein"			
	/codon_start=1			
	/product="At2g43410/T1024.15"			
	/protein_id="FAM19833.1"			

```

/db_xref="GI:20453185"
/translation="MALSMKPFRAIDSGFQSNMLWGLPETTESDLTELEFGYDI
DRITVSSRGEAFIYRHRVEEAAVAAKALOGANLNGSOILEXARPKPKKSLMWGI
GPNYSKDLEEFESKKEKIEDPRLERKPAFIDYMDALQASXNGKPMGSEFLR
VDPLRSAPKREOMAGSYDRNKNMKNKPOYPHYSEFKGDVDSKVLWGLPETATQ
CNDQOILJHNMALIFGEIERKSTPSRNFALVERSAEAEQKEGLOGRLENNPRI
MYSDNLEPPRODSTSEYSGMKRSTDMFNDSFVSPSTGILPSMRPRGTERSEY
NGAEYNDVPEKEPNMRSPSANGTILPSPTGPIIPSAQTRPMRSPNSDSEGYD
AOLVRESKRTRDQSYDGFPMGVDSERFGRSVAARIRKPGDSDHIMRMTAKGSP
PCCARCVPKMGKIEKLEPEVNCASRTDMLAKHYAAVIGCEIYFVFPVREDRPS
VTEERLRSKDRAGVAKLDDGTTLLPVPSPDLPVLQYTRORBLXVYLKLPETAV
PVYASTROESQSNPLHMDARDSPNASHLYPRENTYRGAPEHLTAASKSVSPB
LRIENNAAPQAGVSLPELALATLASILIPASQPAAPESHQPMSPSPSVSTADSNL
YNGEAPQAMKRPQTVHDAASNOSFOOYGNQYPAQGLPEPSPRYPASNNPYTSGP
VHGNMAYOSQSNVMPOLSLPNMHNHNSMTQSSNHHPVOPMVOQOYEASMPDN
YGPILPYOANFHVTTNOMNLNPSQFOAMOPPADKALPEPONALRIQPMISDGC
OGTTDGEVDKNORYSTLQFPAANLLIQIOKKRQOSSGTPAGGP"
misc_difference 2917
/ncore="compared to genome"
/replace="a"
BASE COUNT      983 a      788 c      743 g      984 t
ORIGIN
Query Match      99.8%; Score 2701.2; DB 8; Length 3498;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2703; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGGCGTTATCTATGAAGCCATTCAGAGCCGATGATTCGGTTTCCAGTCAAAACATCTT 60
Db 255 ATGGCGTTATCTATGAAGCCATTCAGAGCCGATGATTCGGTTTCCAGTCAAAACATCTT 314
QY 61 TGGGTGGTACCTTACACCGCCGAGAGACAGACAGTCAAGTCAACCGAGTTGTTGAGA 120
Db 315 TGGGTGGTACCTTACACCGCCGAGAGACAGACAGTCAAGTCAACCGAGTTGTTGAGA 374
QY 121 TAGGCGATATGATAGATACAGGTGTATCTTCCAGAGGCTTGGGTTATATACTAC 180
Db 375 TAGGCGATATGATAGATACAGGTGTATCTTCCAGAGGCTTGGGTTATATACTAC 434
QY 181 AGACATGTGAGAGACAGTGCAGCCAAAGAGCTCTTCAAGAGCAATTTGAATGA 240
Db 435 AGACATGTGAGAGACAGTGCAGCCAAAGAGCTCTTCAAGAGCAATTTGAATGA 494
QY 241 AGTCAATTAAGATTCAGATACGACAGACCGGCAAAACCTTGAAGTCTATGGTGGT 300
Db 495 AGTCAATTAAGATTCAGATACGACAGACCGGCAAAACCTTGAAGTCTATGGTGGT 554
QY 301 GGAATCGGCCCTATGTCTCCAGAGTGAACCTGGAGAAAGTTTCAAGAGTTTGGAAA 360
Db 555 GGAATCGGCCCTATGTCTCCAGAGTGAACCTGGAGAAAGTTTCAAGAGTTTGGAAA 614
QY 361 ATCGAGAGTTTATGTTTCTCAGAGAACCGCAGACAGCTTCAATGATTATATGAGAT 420
Db 615 ATCGAGAGTTTATGTTTCTCAGAGAACCGCAGACAGCTTCAATGATTATATGAGAT 674
QY 421 GATGATGCTTTACAGCTTAAGACATGAAGTGAAGAACCTATGGGTGGTGGCTTTGGCG 480
Db 675 GATGATGCTTTACAGCTTAAGACATGAAGTGAAGAACCTATGGGTGGTGGCTTTGGCG 734
QY 481 GTTGATTTTCTCGGTACAGAGCCCAAAAAAGAACATGGGTGGCTTTACGATTAAC 540
Db 735 GTTGATTTTCTCGGTACAGAGCCCAAAAAAGAACATGGGTGGCTTTACGATTAAC 794
QY 541 AGAAATGGCAATATGATATATAACCGCATATCTCATATGATGAAGACTTTAAAGA 600
Db 795 AGAAATGGCAATATGATATATAACCGCATATCTCATATGATGAAGACTTTAAAGA 854
QY 601 GATGTCAGCAAGTAAGTTCGTGATGTTGGTTCCTCCCTACTACTCTACACAAATGCAAT 660
Db 855 GATGTCAGCAAGTAAGTTCGTGATGTTGGTTCCTCCCTACTACTCTACACAAATGCAAT 914
QY 661 GATGAGCAATATTCGACAAATGCGATGACTCTTGTGATGATGAGAGGTTAAAGAT 720

```

```

Db 915 GATGAGCAATATTCGACAAATGGAGTGAATGACTCTTTGGTGAATCGAGAGGTTAAAGAT 974
QY 721 TACCATCAAGAAATTTTTCGACTTTGGAGTTTGAAGCCCGGAGAGAGCTGCCCAATGC 780
Db 975 TACCATCAAGAAATTTTTCGACTTTGGAGTTTGAAGCCCGGAGAGAGCTGCCCAATGC 1034
QY 781 AAGGAAGCCCTACAGGGAGGATTATCAATAACCTGAAATCAAAATTAATGTTCTCAAC 840
Db 1035 AAGGAAGCCCTACAGGGAGGATTATCAATAACCTGAAATCAAAATTAATGTTCTCAAC 1094
QY 841 GATGATGTCCTCTCGTGAAGCAAGACGATAGTATGTTTACTGTGTATGAAGCGTCAAG 900
Db 1095 GATGATGTCCTCTCGTGAAGCAAGACGATAGTATGTTTACTGTGTATGAAGCGTCAAG 1154
QY 901 ACGATATGTTCAATATGATATGCTTCAATGTTATCTTCCCTATCTACTGAAATTCCT 960
Db 1155 ACGATATGTTCAATATGATATGCTTCAATGTTATCTTCCCTATCTACTGAAATTCCT 1214
QY 961 GGGTCTATGAGGCCCTCAGAGGTAGCAATGACCGTTCAATATATGTCAGAAATCAAT 1020
Db 1215 GGGTCTATGAGGCCCTCAGAGGTAGCAATGACCGTTCAATATATGTCAGAAATCAAT 1274
QY 1021 GACGTTGTTGTTAAGAGCCAAACTGAGAGGCCATCTGCAAAATGGAATGGAATCTC 1080
Db 1275 GACGTTGTTGTTAAGAGCCAAACTGAGAGGCCATCTGCAAAATGGAATGGAATCTC 1334
QY 1081 CCATCTCCACAGAGACCTGGAATCCCTCCATCTCCGACAAAGTACGAGGCCCTATG 1140
Db 1335 CCATCTCCACAGAGACCTGGAATCCCTCCATCTCCGACAAAGTACGAGGCCCTATG 1394
QY 1141 AGGTCAAAACCCGATCTTGGGAAGATATGATCTGCTCAGTTGGTCAAGAAATGAA 1200
Db 1395 AGGTCAAAACCCGATCTTGGGAAGATATGATCTGCTCAGTTGGTCAAGAAATGAA 1454
QY 1201 CGAACCAGAAAGATATGATGATGAGAGGTTTACTCCAAATGGGTGCTGATGAGGCTA 1260
Db 1455 CGAACCAGAAAGATATGATGATGAGAGGTTTACTCCAAATGGGTGCTGATGAGGCTA 1514
QY 1261 TTTGGTGAAGTTCAATGCTGTAGACCTATCCGAGGCCCTGATCTGTACATA 1320
Db 1515 TTTGGTGAAGTTCAATGCTGTAGACCTATCCGAGGCCCTGATCTGTACATA 1574
QY 1321 TGGAGAGGAATGATTCGCAAGAGGTGGAACCTCCGCTGTGTTGCTGTTGTACTATG 1380
Db 1575 TGGAGAGGAATGATTCGCAAGAGGTGGAACCTCCGCTGTGTTGCTGTTGTACTATG 1634
QY 1381 GGAAGGGGATTAACCTAACTGAGGTGAGGTGCAATGTTTCAACAGAACTGATTTG 1440
Db 1635 GGAAGGGGATTAACCTAACTGAGGTGAGGTGCAATGTTTCAACAGAACTGATTTG 1694
QY 1441 AATATGCTGCTAAACATTAAGCCGTTGCCATTTGATGTGAGATCGTTTTTTTCGTACA 1500
Db 1695 AATATGCTGCTAAACATTAAGCCGTTGCCATTTGATGTGAGATCGTTTTTTTCGTACA 1754
QY 1501 GACAGGGAAGAAATTTTGGCTTTACATGAAATTTCTCCGTAACCTTACGCTAAAAGAT 1560
Db 1755 GACAGGGAAGAAATTTTGGCTTTACATGAAATTTCTCCGTAACCTTACGCTAAAAGAT 1814
QY 1561 CGGCGGGGTGTCGCAATTAAGTATGATGATGATCAACTTATCTGTTGGTCCCTACACAGAT 1620
Db 1815 CGGCGGGGTGTCGCAATTAAGTATGATGATGATCAACTTATCTGTTGGTCCCTACACAGAT 1874
QY 1621 TTCTTAATGATGATCTCCAAAGTGAACCGCTTATGATGATGATGATGATGATGATGAT 1680
Db 1875 TTCTTAATGATGATCTCCAAAGTGAACCGCTTATGATGATGATGATGATGATGATGATGAT 1934
QY 1681 TTACCCCGCGCAGCGCTTCTGTTACAGCATATACAGCAAGAAATTCAGTCAATCCT 1740
Db 1935 TTACCCCGCGCAGCGCTTCTGTTACAGCATATACAGCAAGAAATTCAGTCAATCCT 1994
QY 1741 CTGCATTAATGATGATCAAGCCGCGGATTCACCTGCAATGCTAGTACAGTTAATCTCT 1800
Db 1995 CTGCATTAATGATGATCAAGCCGCGGATTCACCTGCAATGCTAGTACAGTTAATCTCT 2054

```

OY	1801	CCCTGAGGAAATTCATTATGGGGGGACACGAAACATTTGACGTGCTCAAAACATCT	1880
Db	2055	CTTGGGGAATTAATTCATTATGGGGGGACACGAAACATTTGACGTGCTCAAAACATCT	2114
OY	1861	GTATGACGACGCTCTCAGAAATACCTAAATAATGACAGCGCTCAAGCTGGGGTATGTTTAACT	1920
Db	2115	GTATGACGACGCTCTCAGAAATACCTAAATAATGACAGCGCTCAAGCTGGGGTATGTTTAACT	2174
OY	1921	CCGGAGCTTTTATGACCTCTGGCAATCTATTCCTGCAACCTTCCAAACCTGCTGCCCT	1980
Db	2175	CCGGAGCTTTTATGACCTCTGGCAATCTATTCCTGCAACCTTCCAAACCTGCTGCCCT	2234
OY	1981	GAGAGTACACCACTATGTGTCAGGACCTTCAACATTTGTTTCCAGACATCACTCCAAAT	2040
Db	2235	GAGAGTACACCACTATGTGTCAGGACCTTCAACATTTGTTTCCAGACATCACTCCAAAT	2294
OY	2041	GGACTGTACAAATGGAGAAGACACCGCTCTCAAGCTTGGAAAAAGAGGTCCACAAACAGTTTCAT	2100
Db	2295	GGACTGTACAAATGGAGAAGACACCGCTCTCAAGCTTGGAAAAAGAGGTCCACAAACAGTTTCAT	2354
OY	2101	GATGCGCTCAAAATCAGTCATTTCCAACAAATAGGGAATCAGTACACTCCAGCTGGGCAACTA	2160
Db	2355	GATGCGCTCAAAATCAGTCATTTCCAACAAATAGGGAATCAGTACACTCCAGCTGGGCAACTA	2414
OY	2161	CCCTCCCTCCCTCTGGGCTTATGCCCTCCAGGCTTCAACAAACCCCACTACACTAGTGGAAATG	2220
Db	2415	CCCTCCCTCCCTCTGGGCTTATGCCCTCCAGGCTTCAACAAACCCCACTACACTAGTGGAAATG	2474
OY	2221	GTCCATGGGCAACATGACATAACACAGAGCAATCTGTAAACATGCGCTCAGCTGTCCGCTTA	2280
Db	2475	GTCCATGGGCAACATGACATAACACAGAGCAATCTGTAAACATGCGCTCAGCTGTCCGCTTA	2534
OY	2281	CCAAATATGCTCTCAATTAATTAATTATTCATGTACACTCAGGGTTCGTCAATCATCTGTT	2340
Db	2535	CCAAATATGCTCTCAATTAATTAATTATTCATGTACACTCAGGGTTCGTCAATCATCTGTT	2594
OY	2341	TTCTAGGCGCAATGGTGCACGCAATACCACACGAAAGCGTCCATGCCAAACCCAAACATATAGT	2400
Db	2595	TTCTAGGCGCAATGGTGCACGCAATACCACACGAAAGCGTCCATGCCAAACCCAAACATATAGT	2654
OY	2401	CCAAATTTCCAGTTATTCAGCAAGCTAAATTTTTCATGGCTTAACAAACAAATTCAGGACAGAAC	2460
Db	2655	CCAAATTTCCAGTTATTCAGCAAGCTAAATTTTTCATGGCTTAACAAACAAATTCAGGACAGAAC	2714
OY	2461	TTAAACCCCTTCCCAATTTCAAGCTGCATGCAACACACAGCAGATAAGGCAAAATTTAGAG	2520
Db	2715	TTAAACCCCTTCCCAATTTCAAGCTGCATGCAACACACAGCAGATAAGGCAAAATTTAGAG	2774
OY	2521	CCAAACAAACCAAGACATACCATTTGCAACCGCTATGATCTCTGGGGATTTGGTCAGGGTACACAA	2580
Db	2775	CCAAACAAACCAAGACATACCATTTGCAACCGCTATGATCTCTGGGGATTTGGTCAGGGTACACAA	2834
OY	2581	GATGGGAGAGTGCATTAAGAAATACAGAGATACAGTCAACACATCAATTAATTGGACCAACCTT	2640
Db	2835	GATGGGAGAGTGCATTAAGAAATACAGAGATACAGTCAACACATCAATTAATTGGACCAACCTT	2894
OY	2641	CTTCTCCAGATACAGCAGAAACACAGCAACAGTCTTCAGGTACTCCGGCTGACAGAGGG	2700
Db	2895	CTTCTCCAGATACAGCAGAAACACAGCAACAGTCTTCAGGTACTCCGGCTGACAGAGGG	2954
OY	2701	CCCTTGA 2706	
Db	2955	CCCTTGA 2960	

[illegible]

SOURCE	thale cress.
ORGANISM	Arabidopsis thaliana
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
AUTHORS	1 Amadio, R.M., Schomburg, F.M., Michaels, S.D. and Patton, D.
TITLE	Floral induction gene
JOURNAL	Patent: WO 0212518-A 1 14-FEB-2002;
FEATURES	WISCONSIN ALUMNI RESEARCH FOUNDATION (US)
source	Location/Qualifiers 1..4593 /organism="Arabidopsis thaliana" /db_xref="taxon:3702"
BASE COUNT	1275 a 958 c 978 g 1382 t
ORIGIN	
Query Match	70.1%; Score 1896.8; DB 6; Length 4593;
Best Local Similarity	90.2%; Pred. No. 0;
Matches 2137; Conservative	0; Mismatches 2; Indels 229; Gaps 1;
QY	568 CAGTATCTCTCACTCATATGAAAGACTTTTAAAGAGATGTCCAGCCCAAGTAAGTCTGTGCG 627
DB	2226 CAGTATCTCTCACTCATATGAAAGACTTTTAAAGAGATGTCCAGCCCAAGTAAGTCTGTGCG 2285
OY	628 ATTGGGGTCCCTCCACTGCTACACAAATGCAATGATGTGACAAATTCTGCACATGCGATG 687
DB	2286 ATTGGGGTCCCTCCACTGCTACACAAATGCAATGATGTGACAAATTCTGCACATGCGATG 2345
OY	688 ATACTCTTTGGTGAGATCGAGAGGGGTAAAAAGTTACCATCAAGGAATTTTGCACCTTGTG 747
DB	2346 ATACTCTTTGGTGAGATCGAGAGGGGTAAAAAGTTACCATCAAGGAATTTTGCACCTTGTG 2405
OY	748 GAGTTTAGACCGGGAGAGAGCTCCCAATGCAAGAAAGGCCCTACAGGGGAGGTTATTC 807
DB	2406 GAGTTTAGACCGGGAGAGAGCTCCCAATGCAAGAAAGGCCCTACAGGGGAGGTTATTC 2465
OY	808 AATATCCTCAATATCAAAATTTATGTACTCAACGATGTGAGTTCCTCCCTGACCAAGACGAT 867
DB	2466 AATATCCTCAATATCAAAATTTATGTACTCAACGATGTGAGTTCCTCCCTGACCAAGACGAT 2525
OY	868 ACTAGTTTTTACTCTGTGTATGAAAACGGTCAAGACAGATATGTTCAATATGATCTCTCA 927
DB	2526 ACTAGTTTTTACTCTGTGTATGAAAACGGTCAAGACAGATATGTTCAATATGATCTCTCA 2585
OY	928 TGTGATCTCTCTCTCATCTTACTAGGGAATTCCTGGGCTATGAGGCCCCCTAGAGGTACG 987
DB	2586 TGTGATCTCTCTCTCATCTTACTAGGGAATTCCTGGGCTATGAGGCCCCCTAGAGGTACG 2645
OY	988 AATGAGCGTTCATTAATGATGTGCAGAAATACATGACGTTTGGTAAAGAGCCCAAACTCG 1047
DB	2646 AATGAGCGTTCATTAATGATGTGCAGAAATACATGACGTTTGGTAAAGAGCCCAAACTCG 2705
OY	1048 AGAAGGCCATGTGCAAATGGAATCTGGAATTAATCCCATCTCCAACAGACCTGGAATCTCT 1107
DB	2706 AGAAGGCCATGTGCAAATGGAATTAATCCCATCTCCAACAGACCTGGAATCTCT 2765
OY	1108 CCATCTCCCTGACACAGATGACGAGGGGCCCTATGAGGTGAACCCCGATCTTTGGGAAGGA 1167
DB	2766 CCATCTCCCTGACACAGATGACGAGGGGCCCTATGAGGTGAACCCCGATCTTTGGGAAGGA 2825
OY	1168 TATGATCCTGCTCAGTTGGTCAGAGAAAAGTAAAGCAACAGAAAGATGATCAGTGGAC 1227
DB	2826 TATGATCCTGCTCAGTTGGTCAGAGAAAAGTAAAGCAACAGAAAGATGATCAGTGGAC 2885
OY	1228 GGTATTACTCCAAATGGGTGTGATGACATATGAGAGAGAAAGATGGCCCAAGGGTGA 1287
DB	2886 GGTATTACTCCAAATGGGTGTGATGACATATGAGAGAGAAAGATGGCCCAAGGGTGA 2945
OY	1288 CCTATCCGTGGCCCCCCTGATTTGATGACATATGAGAGAGAAAGATGGCCCAAGGGTGA 1347
DB	2946 CCTATCCGTGGCCCCCCTGATTTGATGACATATGAGAGAGAAAGATGGCCCAAGGGTGA 3005

QY 1348 ACTCCGCTCTGTTGCTGCTGTTGTACTATATGGGAAAGGAGTTGAACATACT----- 1403
 |||||
 Db 3006 ACTCCGCTCTGTTGCTGCTGTTGTACTATATGGGAAAGGAGTTGAACATACTGTGA 3005
 |||||
 QY 1404 ----- 1403
 Db 3066 GTACTAATTTCTAGCACTTTAACCCTTCTAGTGTCTTTCTTTTCAGAGCATTTATATAT 3125
 |||||
 QY 1404 ----- 1403
 Db 3126 TTTCATTTGATTTCTGATGAGATGCTTGATTTTCATGTTGATTTGATTTTGGCATTTG 3185
 |||||
 QY 1404 ----- 1403
 Db 3186 TTACTGTTAGTTTCTGAGATGCTTGATTTTCATGTTGATTTGATTTTGGCATTTG 3245
 |||||
 QY 1404 ----- 1403
 Db 3246 CCTCAATTAAGTACTGTTGTTTCTTTTATTAATGATTTATAGGCTGAGGCTGCA 3305
 |||||
 QY 1419 TTGTTACAGCAAGAACTGATTTGAATATGCTGCTAAACATTTAGCGGCTTGCAATTTGATG 1478
 |||||
 Db 3306 TTGTTACAGCAAGAACTGATTTGAATATGCTGCTAAACATTTAGCGGCTTGCAATTTGATG 3365
 |||||
 QY 1479 TGAGATCGTTTCTGTTGCTACCAAGAGGAGAAAGATTTTGGCTTTACACTGAATTTCT 1538
 |||||
 Db 3366 TGAGATCGTTTCTGTTGCTACCAAGAGGAGAAAGATTTTGGCTTTACACTGAATTTCT 3425
 |||||
 QY 1539 CCGGTACCTTAGGCTCAAAAGATGGGGGGGTTGGCAATTTAGATGATGTTGCAACTTT 1598
 |||||
 Db 3426 CCGGTACCTTAGGCTCAAAAGATGGGGGGGTTGGCAATTTAGATGATGTTGCAACTTT 3485
 |||||
 QY 1599 ATTCTGTTGCTCCATCAGATTTCTTACTGATGATGATGATGATGATGATGATGATGATG 1658
 |||||
 Db 3486 ATTCTGTTGCTCCATCAGATTTCTTACTGATGATGATGATGATGATGATGATGATGATG 3545
 |||||
 QY 1659 GCTAATNGTGTGTTCTCTCAAGTTACCCCGGCAAGCGGCTGCTTTACAGATCATACAG 1718
 |||||
 Db 3546 GCTAATNGTGTGTTCTCTCAAGTTACCCCGGCAAGCGGCTGCTTTACAGATCATACAG 3605
 |||||
 QY 1719 ACAAGATCTAGTCCATCTGCTGATTTATGATGATGATGATGATGATGATGATGATGATG 1778
 |||||
 Db 3606 ACAAGATCTAGTCCATCTGCTGATTTATGATGATGATGATGATGATGATGATGATGATG 3665
 |||||
 QY 1779 TGCTAGTACAGTTTATATCTCTCTAGGAAATTAACATTTAGGGGTGACACAGATTT 1838
 |||||
 Db 3666 TGCTAGTACAGTTTATATCTCTCTAGGAAATTAACATTTAGGGGTGACACAGATTT 3725
 |||||
 QY 1839 GACAGTGTCTTCAAAACATCTGTGTAGCGAGCTCTCAGATACCTAATTAATGACAGCGC 1898
 |||||
 Db 3726 GACAGTGTCTTCAAAACATCTGTGTAGCGAGCTCTCAGATACCTAATTAATGACAGCGC 3785
 |||||
 QY 1899 TCAAGTGGGGTATGTTTAACTCCGAGAGCTTTTGGCAGCTGTCATATTTCTCCCTGC 1958
 |||||
 Db 3786 TCAAGTGGGGTATGTTTAACTCCGAGAGCTTTTGGCAGCTGTCATATTTCTCCCTGC 3845
 |||||
 QY 1959 AACTTTTCAACCTGCTGCCCCCTAGAGTACCACTATGTCAGACCTTCAACAGTTGT 2018
 |||||
 Db 3846 AACTTTTCAACCTGCTGCCCCCTAGAGTACCACTATGTCAGACCTTCAACAGTTGT 3905
 |||||
 QY 2019 TTTCACAGCAGATCAGTCCAAATGAGTATGATGATGATGATGATGATGATGATGATGATG 2078
 |||||
 Db 3906 TTTCACAGCAGATCAGTCCAAATGAGTATGATGATGATGATGATGATGATGATGATGATG 3965
 |||||
 QY 2079 AAGAGTTCACAAACAGTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2138
 |||||
 Db 3966 AAGAGTTCACAAACAGTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4025
 |||||
 QY 2139 GTTACATCTCAGCTGGGCAACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2198
 |||||
 Db 4026 GTTACATCTCAGCTGGGCAACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4085
 |||||
 QY 2199 CCCCACTACACTAGTGAATGTTCCATGAGCAACATGCAATACAGAGCATCTCTTAA 2258
 |||||

Db 4086 CCCCACTACACTAGTGAATGTTCCATGAGCAACATGCAATGATGATGATGATGATGATGATG 4145
 |||||
 QY 2259 CATGCTCAGCTGCTCTCCGTTTACCAATATGCTCATATATATTTATTCATGATGATGATGATG 2318
 |||||
 Db 4146 CATGCTCAGCTGCTCTCCGTTTACCAATATGCTCATATATATTTATTCATGATGATGATGATG 4205
 |||||
 QY 2319 GGGTGTGCAATCATCT 2378
 |||||
 Db 4206 GGGTGTGCAATCATCT 4265
 |||||
 QY 2379 CATGCCAAACCAAAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2438
 |||||
 Db 4266 CATGCCAAACCAAAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4325
 |||||
 QY 2439 AACACAAATCAGGACAGATTTAAACCTTCCCAATTTCAAGTGCATGCAACACAC 2498
 |||||
 Db 4326 AACACAAATCAGGACAGATTTAAACCTTCCCAATTTCAAGTGCATGCAACACAC 4385
 |||||
 QY 2499 AGCAGATTAAGCAATTTTACAGCACAACCAAGCACTAGATGATGATGATGATGATGATG 2558
 |||||
 Db 4386 AGCAGATTAAGCAATTTTACAGCACAACCAAGCACTAGATGATGATGATGATGATGATG 4445
 |||||
 QY 2559 TGGGGATGCTCAGGATACCAACAGATGGGAGTGTGATTAAGATCAGAGATACAGTCAAC 2618
 |||||
 Db 4446 TGGGGATGCTCAGGATACCAACAGATGGGAGTGTGATTAAGATCAGAGATACAGTCAAC 4505
 |||||
 QY 2619 ACTACATTTTGACAGCAACCTTTCTTCCAGATACAGCAGAAACACAGACAGATCTTTC 2678
 |||||
 Db 4506 ACTACATTTTGACAGCAACCTTTCTTCCAGATACAGCAGAAACACAGACAGATCTTTC 4565
 |||||
 QY 2679 AGGTACTCCGGCTGACAGAGGGGCTTGA 2706
 |||||
 Db 4566 AGGTACTCCGGCTGACAGAGGGGCTTGA 4593
 |||||

RESULT 4
 AC002335/c 102057 bp DNA linear PLIN 11-MAR-2002
 LOCUS Arabidopsis thaliana chromosome 2 clone T1024 map C1C10F02,
 DEFINITION complete sequence.
 ACCESSION AC002335
 VERSION AC002335.3 GI:20196857
 KEYWORDS HTG.
 SOURCE Arabidopsis thaliana.
 ORGANISM Arabidopsis thaliana.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eustosids II; Brassicales; Brassicaceae; Arabidopsis.
 REFERENCE 1 (bases 1 to 102057)
 Rounsley,S.D., Teschduy,M.M., Lin,X., Ketchum,K.A., Crosby,M.L.,
 Brandon,R.C., Spriggs,T.A., Mason,T.M., Kerlavage,A.R., Adams,M.D.,
 Somerville,C.R. and Venter,J.C.
 Somerville,C.R. and Venter,J.C.
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 102057)
 Lin,X.
 JOURNAL Direct Submission
 TITLE Submitted (09-MAR-2000) The Institute for Genomic Research, 9712
 Medical Center Dr., Rockville, MD 20850, USA
 JOURNAL 3 (bases 1 to 102057)
 Town,C.D. and Kaul,S.
 JOURNAL Direct Submission
 TITLE Submitted (27-FEB-2002) The Institute for Genomic Research, 9712
 Medical Center Dr., Rockville, MD 20850, USA, cdtown@tigr.org
 COMMENT On Apr 18, 2002 this sequence version replaced gi:6598343.
 FEATURES
 source
 1..102057
 /organism="Arabidopsis thaliana"
 /cultivar="Columbia"
 /db_xref="taxon:3702"
 /chromosome="2"
 /map="C1C10F02"
 /clone="T1024"

```

misc_feature      1..16126
                  /note="overlap with BAC clone F14B2
                  (AC004430:74158..90283)."
```

repeat_region

```

gene             /complement(334..335)
                  /rpt_family="AT_rich"
                  /complement(3113..4397)
                  /gene="At2g43270"
                  /note="T1024.1"
```

mRNA

```

                  /complement(join(<3113..3691,3734..3924,4004..>4397))
                  /gene="At2g43270"
                  /complement(join(3113..3691,3734..3924,4004..4397))
                  /gene="At2g43270"
                  /codon_start=1
                  /product="putative myb-related transcription factor"
                  /protein_id="AAB64309.2"
                  /db_xref="GI:20196861"
```

CDS

```

                  /translation="MIVVVDLLEIEIFLGPLKSLNFKTVSKQMSILSKSPEARR
                  LNVKKEKILAVDRELFEGEEIEQMYLHCDIDATRSLCEGLVCIAPGMIVY
                  LNPSTROLRRFPSPNHHVPSDRIRFOREFELTSPGNMKGFGDRKNGSVYVR
                  MCFSPVEKCEVLDERGWESELNPPRIDVGRKSRPGIILDLHTLEPNNVSPSC
                  CTMSAQVNLDRLAITKTKNGAPWEWEIISLDVDKREITSYITSLAIIINHPPE
                  SSWFVLTYSKLNLFVCDNKKRLFKYNTETNELCCSDISVISPLEDLVPLRDS
                  GHDRDYKIRKTSSCGLESKHPEAGLEFKEIITFRFVATLVGVQYRSLSKFK"
                  5148..5767
                  /gene="At2g43280"
                  /note="T1024.2; Predicted by genscan"
                  /join(<5148..5447,5552..>5767)
                  /gene="At2g43280"
                  /join(5148..5447,5552..5767)
                  /gene="At2g43280"
                  /codon_start=1
                  /product="hypothetical protein"
                  /protein_id="AAB64334.1"
                  /db_xref="GI:2289005"
                  /translation="MFYDDYSRRIGFVWRVMSCRSEKDRILARRCCKEHCQVSI
                  RGRGFSYRKRPRSTRECKAMIHVKYDRSGKWTYTKRKHNPGLVYSPPEARLTLE
                  KDRIOELTELENNKRLCAAYKEQDAFAKIVYEHSNQIAKVENYVNNLKEEHEHLE
                  HALRSKDAI"
```

mRNA

```

                  /complement(5982..6953)
                  /gene="At2g43290"
                  /note="T1024.3; supported by full length cDNA:
                  Ceres:31535"
                  /complement(<5982..>6953)
                  /gene="At2g43290"
                  /complement(6214..6861)
                  /gene="At2g43290"
                  /codon_start=1
                  /product="putative calcium binding protein"
                  /protein_id="AAB64310.2"
                  /db_xref="GI:20196862"
```

gene

```

                  /translation="WVRILYINILNSFLSLVPKKLTLPPLSPWPKTLHKNSPSP
                  STMLPSPSSSAPTRKIDPSLKRVMFQMPKNGGRITKELNLSNLTGIYIPDKL
                  TOMIHKIDANGDCVDIDEFESLSSIVDEHNDGFEEDMKDAFVFPDGDGFFIT
                  VEELKSYMASLGGKOGTKLDGCKKIMQVADGDGRVNYKEFLQMMKGGSSSSN"
```

repeat_region

```

repeat_region    /complement(6998..7029)
                  /rpt_family="AT_rich"
                  /complement(7486..7547)
                  /rpt_family="AT_rich"
                  /complement(7628..7660)
                  /rpt_family="AT_rich"
                  /complement(10479..10550)
                  /gene="At2g43300"
                  /complement(10479..10550)
                  /gene="At2g43300"
                  /complement(10479..10550)
                  /gene="At2g43300"
                  /product="tRNA-Thr"
                  11012..11413
                  /gene="At2g43310"
                  /note="T1024.5; predicted by genscan"
                  <11012..>11413
                  /gene="At2g43310"
                  11012..11413
                  /gene="At2g43310"
                  /codon_start=1
```

gene

```

                  /product="hypothetical protein"
                  /protein_id="AAB64333.1"
                  /db_xref="GI:2289004"
                  /translation="MNGTNRHRLQLVLSCRKTYAOTQPGSSTIIMASSSDEFLA
                  KSRANLRRPFSNHFMDSKTASRGEKLGRLGADVVSIDADEEISRIHRRKV
                  LPLEDSVRRRTGIRVDGTQELNDISVPPIPNK"
                  /complement(11649..14833)
                  /gene="At2g43320"
                  /note="T1024.6; predicted by genscan; supported by cDNA:
                  g1_14517475_db_AY039573.1"
                  /complement(join(<11649..11950,12040..12519,12610..12654,
                  13002..13102,13191..13257,13742..13804,13936..14107,
                  14566..14833))
                  /gene="At2g43320"
                  /complement(join(11780..11950,12040..12519,12610..12654,
                  13002..13102,13191..13257,13742..13804,13936..14064))
                  /gene="At2g43320"
                  /codon_start=1
                  /product="expressed protein"
                  /protein_id="AAB64311.2"
                  /db_xref="GI:20196858"
                  /translation="WRAPSLAQCPLPDLPODRGCVALSSEKIDLPPTPAVEIPST
                  VAHRRYSGENIDALGLOVFKGVSVADIIIGLSETAPLNKESLKSWESSVAVNLV
                  KNEIRDGLSFRGKRVLELCNFGVPGIIFACLGASSVHFDLSAETIRCTIPNVLA
                  NLEQARDROSROPESPLTPSRQATASAVRFAGWEELSTVLSIRFDLESPISAPNN
                  LSPSEDFMDGCSODGSTIGDPDPSRSRSKLSGSAMERANETGGEGYGVYIIM
                  TEIPYSITSLKLKLSLRKCLRPPYGVYLAARKQYGFNSGAKHLNLDDEITLGA
                  HLVEETDRDIWKEFLK"
                  16041..18948
                  /gene="At2g43330"
                  /note="T1024.7"
                  /join(<16041..16188,16408..16514,16563..16898,17700..17993,
                  18080..18948)
                  /gene="At2g43330"
                  /join(16041..16188,16408..16514,16563..16898,17700..17993,
                  18080..18760)
                  /gene="At2g43330"
                  /codon_start=1
                  /product="putative membrane transporter"
                  /protein_id="AAB64332.1"
                  /db_xref="GI:2289003"
                  /translation="WTLTPNAPGSSGYLDMFPRRRSYFGNSYILGIVYAGIGGL
                  FGVDYVIGSGLILYIKDDEYVROSSFOYVNSFSFSSKLETTIVSMALVGMAGAA
                  GGMTNDVYGRKATLPADVFAAGAIYMAAPADPVYVLSIGLIGLVAGVAVAPAY
                  IAEASPESEVKGIVSTVNLMTGQFLISLVNKAFTQVPGYRMMILSYSGVPAYQRT
                  LMLFPPSPRWLFEMKRRKAELIQVLAFTYDLSLELDLHLSAEEEKOKKRVGTL
                  DVFRSKELRLAFLAGAGIQAOQFTGINTVYVYPTIYQVAGFSNQLAFLSLIVA
                  MNAAGTYVGIYFIDHCRKLLALSILFGVILSLILSVSEFKQSETSDDGLYGMIAV
                  LGLALTYVFPAGMGPVPMVYSEIYPOOYIGCGSAYVNTSNLIVAQTFITIE
                  AAGCGMFLIAGIYAVLAVIFYVYFVETQGLTSEVQIKERAYGNISMGSSSS
                  NMKGLEBOGSSQ"
                  /complement(21849..21900)
                  /rpt_family="AT_rich"
                  22675..23510
                  /gene="At2g43340"
                  /note="T1024.8"
                  /join(<22675..23164,23243..23304,23325..23510)
                  /gene="At2g43340"
                  /join(22675..23164,23243..23304,23325..23498)
                  /gene="At2g43340"
                  /codon_start=1
                  /product="unknown protein"
                  /protein_id="AAB64331.1"
                  /db_xref="GI:2289002"
                  /translation="MCDSDFLSRCSYEADVDSDEISTSSFSQSYSESEERINNG
                  FGVYOSKQTKLEKKSNVLEGGYVDSANDDLKTRKSLTDLDLELKGCVDLGFG
                  FNYEIEPELQNTLPALCEYSMQKFIIDODHNNHSSPEKSSVLDSPVPIASMTI
                  SSPGDNPDDVYARLKFMAQVAYACTMRRGCFDQPKABESMEBEDEDTQFAGVSKLML
                  TCLKTTIVELFAHTVCIKRLMT"
                  /complement(23549..25505)
                  /gene="At2g43350"
                  /note="T1024.9; supported by full length cDNA:
                  Ceres:30737"
```


[illegible]

Db 41870 GTACACTCCAGCTGGGCAACTACGCTCTCTCGCTTGGGTTACCCCTCCAGGTTCAACAA 41811
Qy 2199 CCCCAACTACACTAGTGGAAATGCTCATGCGCAACATGCAATACAGAGCCAACTCTTAA 2258
Db 41810 CCCCAACTACACTAGTGGAAATGCTCATGCGCAACATGCAATACAGAGCCAACTCTTAA 41751
Qy 2259 CATGCTCAGCTGATCTCGCTTACCAAAATATGCTCATATATATTTTCCATGATACACA 2318
Db 41750 CATGCTCAGCTGATCTCGCTTACCAAAATATGCTCATATATATTTTCCATGATACACA 41651
Qy 2319 GGGTTCTGCAATATCATCTGTTTCTCAGCCCATAGGTCAGCAATACCAACGAAAGCGTC 2378
Db 41690 GGGTTCTGCAATATCATCTGTTTCTCAGCCCATAGGTCAGCAATACCAACGAAAGCGTC 41631
Qy 2379 CATGCCAAACCAAACTATGCTCCAAATTCGAAGTTATACGAAGCTAATTTTTCATGGCGT 2438
Db 41630 CATGCCAAACCAAACTATGCTCCAAATTCGAAGTTATACGAAGCTAATTTTTCATGGCGT 41571
Qy 2439 AACACAAATCAGGACAGCAACTTAAACCTTCCCAATTTTCAAGCTGCATGCAACGAC 2498
Db 41570 AACACAAATCAGGACAGCAACTTAAACCTTCCCAATTTTCAAGCTGCATGCAACGAC 41511
Qy 2499 AGCAGATTAAGCAATTTTAAAGCCCAAAACCAAGCACTACGATTGCAAGCCATGATCTC 2558
Db 41510 AGCAGATTAAGCAATTTTAAAGCCCAAAACCAAGCACTACGATTGCAAGCCATGATCTC 41451
Qy 2559 TGGGATGCTGATGCTGATACAGATGGGAGGTGATTAAGATTCAGATACAGATCAAC 2618
Db 41450 TGGGATGCTGATGCTGATACAGATGGGAGGTGATTAAGATTCAGATACAGATCAAC 41391
Qy 2619 ACTCAATTTGACGAAACCTTCTTCTCAGATACAGAGAAACGACGCAACAGCTCTTC 2678
Db 41390 ACTCAATTTGACGAAACCTTCTTCTCAGATACAGAGAAACGACGCAACAGCTCTTC 41331
Qy 2679 AGGTACTCCGGCTGGACAGGGGCGCTTGA 2706
Db 41330 AGGTACTCCGGCTGGACAGGGGCGCTTGA 41303

RESULT 5
AX420422/c 801 bp DNA linear PAT 18-JUN-2002
LOCUS AX420422
DEFINITION Sequence 4 from Patent WO0212518.
ACCESSION AX420422
VERSION AX420422.1 GI:21524578
KEYWORDS
SOURCE
ORGANISM
thale cress.
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE
1 Amasino, R.M., Schomburg, F.M., Michaels, S.D. and Patton, D.
TITLE Floral induction gene
JOURNAL Patent: WO 0212518-A 4 14-FEB-2002;
WISCONSIN ALUMNI RESEARCH FOUNDATION (US)
FEATURES
Source
Location/Qualifiers
1..801
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/note="FPA antisense fragment"

BASE COUNT 206 a 207 c 166 g 222 t
ORIGIN
Query Match 29.5%; Score 797.8; DB 6; Length 801;
Best Local Similarity 99.8%; Pred. No. 2.1e-220;
Matches 799; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 587 AAGACTTTAAAGAGATGTCAGCCAAATAGGTTCTGTGGATTTGGTTCCCTCTACTG 646
Db 801 AAGACTTTAAAGAGATGTCAGCCAAATAGGTTCTGTGGATTTGGTTCCCTCTACTG 742
Qy 647 CTACACATGCAATGATGAGCAAAATTTGCACAATGATGATCTTTGGTGATGCG 706

Db 741 CTACACATGCAATGATGAGCAAAATTTGCACAATGATGATCTTTGGTGATGCG 682
Qy 707 AGAGGGTAAAAAGTTTACCATCAAGAAATTTTGCACCTGTGTGAGATTAGAGCGCGAGG 766
Db 681 AGAGGGTAAAAAGTTTACCATCAAGAAATTTTGCACCTGTGTGAGATTAGAGCGCGAGG 622
Qy 767 AAGCTCGCCATGCAAGGAAGGCTCAGAGGGAGGTTATTCATTAATTCATTAATTCATTA 826
Db 621 AAGCTCGCCATGCAAGGAAGGCTCAGAGGGAGGTTATTCATTAATTCATTAATTCATTA 562
Qy 827 TTATGTACTCAACATGATGCTCTCCGACCAAGCATATCTAGTTTCTCTGTA 886
Db 561 TTATGTACTCAACATGATGCTCTCCGACCAAGCATATCTAGTTTCTCTGTA 502
Qy 887 TGAACGGTCAAGAGACATATGTCATATATGATCTTTCATGATGATCTCTCTCAT 946
Db 501 TGAACGGTCAAGAGACATATGTCATATATGATCTTTCATGATGATCTCTCTCAT 442
Qy 947 CTACTGGAATTCCTGGGCTATGAGGCCCTCAGAGGTACGAATGAGGCTCATATTAATG 1006
Db 441 CTACTGGAATTCCTGGGCTATGAGGCCCTCAGAGGTACGAATGAGGCTCATATTAATG 382
Qy 1007 GTGCAATATACATGACGTTGTTGGTAAAGAGCCAAACGAGAGGCCATCTGCAATG 1066
Db 381 GTGCAATATACATGACGTTGTTGGTAAAGAGCCAAACGAGAGGCCATCTGCAATG 322
Qy 1067 GAATGGAATATCTCCATCTCCACAGAGCACTGGAATCTCCATCTCTGACACAAGTA 1126
Db 321 GAATGGAATATCTCCATCTCCACAGAGCACTGGAATCTCCATCTCTGACACAAGTA 262
Qy 1127 CGAGCGCCCTATGAGGTCAAAACCCGATTTCTGGGAAGATATGATCTGCTCACTTG 1186
Db 261 CGAGCGCCCTATGAGGTCAAAACCCGATTTCTGGGAAGATATGATCTGCTCACTTG 202
Qy 1187 TCAGAGAAATGTAAGCAACCAAGAGATGATGATGATGATGATGATGATGATGATG 1246
Db 201 TCAGAGAAATGTAAGCAACCAAGAGATGATGATGATGATGATGATGATGATGATG 142
Qy 1247 TCGATGAGAGTCAATTTGGTCTGAGGTCTGATGCTGATGACCTATCCGTGCCCCCTG 1306
Db 141 TCGATGAGAGTCAATTTGGTCTGAGGTCTGATGCTGATGACCTATCCGTGCCCCCTG 82
Qy 1307 ATTCTGATCACAATGAGAGGAATGATGATGATGATGATGATGATGATGATGATG 1366
Db 81 ATTCTGATCACAATGAGAGGAATGATGATGATGATGATGATGATGATGATGATG 22
Qy 1367 GTTGTACTATGGAAGG 1387
Db 21 GTTGTACTATGGAAGG 1

RESULT 6
AX420423 349 bp DNA linear PAT 18-JUN-2002
LOCUS AX420423
DEFINITION Sequence 5 from Patent WO0212518.
ACCESSION AX420423
VERSION AX420423.1 GI:21524579
KEYWORDS
SOURCE
ORGANISM
thale cress.
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE
1 Amasino, R.M., Schomburg, F.M., Michaels, S.D. and Patton, D.
TITLE Floral induction gene
JOURNAL Patent: WO 0212518-A 5 14-FEB-2002;
WISCONSIN ALUMNI RESEARCH FOUNDATION (US)
FEATURES
Source
Location/Qualifiers
1..349
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"

BASE COUNT 92 a 68 c 81 g 108 t
 ORIGIN

Query Match 10.0%; Score 271.4; DB 6; Length 349;
 Best Local Similarity 94.6%; Pred. No. 3.1e-67;
 Matches 281; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 1 ATGGCGTTATCTATGAAGCCATTGACAGCCGATGATTCGGTTCCAGTCAAAACATCTT 60
 Db 1 ATGGCGTTATCTATGAAGCCATTGACAGCCGATGATTCGGTTCCAGTCAAAACATCTT 60

QY 61 TGGGTGGTACGCTTAACCGCGAGACAGAGTCAATCATGACCGAGTTGTTGAAGA 120
 Db 61 TGGGTGGTACGCTTAACCGCGAGACAGAGTCAATCATGACCGAGTTGTTGAAGA 120

QY 121 TACGGCGATATGTAATGATACAGCGTATTTCTTCACGAGGCTTGGCTTAATATCTAC 180
 Db 121 TACGGCGATATGTAATGATACAGCGTATTTCTTCACGAGGCTTGGCTTAATATCTAC 180

QY 181 AGACATGTGAGAGACAGTGCAGCCAAAGAGGCTTTCAAGAGCAATTTGAATGA 240
 Db 181 AGACATGTGAGAGAGAGTGCAGCCAAAGAGGCTTTCAAGAGCAATTTGAATGA 240

QY 241 AGTCAATTTAAGATCGAATACGACGACCGGCTTTGTTCTTATCTATCTTCGTTG 297
 Db 241 AGTCAATTTAAGATCGAATACGACGACCGGCTTTGTTCTTATCTATCTTCGTTG 297

RESULT 7
 AX420424 3715 bp DNA linear PAT 18-JUN-2002
 LOCUS
 DEFINITION Sequence 6 from Patent WO0212518.
 ACCESSION AX420424
 VERSION AX420424.1 GI:21524580
 KEYWORDS
 SOURCE
 ORGANISM
 Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 1
 Amasino, R.M., Schomburg, F.M., Michaels, S.D. and Patton, D.
 JOURNAL
 TITLE
 REFERENCE
 AUTHORS
 PATENT: WO 0212518-A 6 14-FEB-2002;
 WISCONSIN ALUMNI RESEARCH FOUNDATION (US)
 FEATURES
 source
 1. 3715
 /organism="Arabidopsis thaliana"
 /db_xref="taxon:3702"
 /note="FPA promoter plus intron"
 BASE COUNT 1055 a 728 c 711 g 1221 t
 ORIGIN

Query Match 10.0%; Score 271.4; DB 6; Length 3715;
 Best Local Similarity 94.6%; Pred. No. 3.3e-67;
 Matches 281; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 1 ATGGCGTTATCTATGAAGCCATTGACAGCCGATGATTCGGTTCCAGTCAAAACATCTT 60
 Db 1832 ATGGCGTTATCTATGAAGCCATTGACAGCCGATGATTCGGTTCCAGTCAAAACATCTT 1891

QY 61 TGGGTGGTACGCTTAACCGCGAGACAGAGTCAATCATGACCGAGTTGTTGAAGA 120
 Db 1892 TGGGTGGTACGCTTAACCGCGAGACAGAGTCAATCATGACCGAGTTGTTGAAGA 1951

QY 121 TACGGCGATATGTAATGATACAGCGTATTTCTTCACGAGGCTTGGCTTAATATCTAC 180
 Db 1952 TACGGCGATATGTAATGATACAGCGTATTTCTTCACGAGGCTTGGCTTAATATCTAC 2011

QY 181 AGACATGTGAGAGACAGTGCAGCCAAAGAGGCTTTCAAGAGCAATTTGAATGA 240
 Db 2012 AGACATGTGAGAGACAGTGCAGCCAAAGAGGCTTTCAAGAGCAATTTGAATGA 2071

QY 241 AGTCAATTTAAGATCGAATACGACGACCGGCAAAACCTTGAAGATCTATGGGTG 297
 Db 2072 AGTCAATTTAAGATCGAATACGACGACCGGTTGTTCTTATCTATCTTCGTTG 2128

RESULT 8
 AC108755/c 154345 bp DNA linear HTG 31-JAN-2002
 LOCUS
 DEFINITION Oryza sativa (japonica cultivar-group) chromosome 9 clone
 OSJNB40028C22, *** SEQUENCING IN PROGRESS ***, 2 ordered pieces.
 AC108755
 AC108755.1 GI:18449960
 HTG; HTGS_PHASE2.
 ORIGIN
 Oryza sativa (japonica cultivar-group).
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzaceae; Oryza.
 1 (bases 1 to 154345)
 Yoon, U.-H., Hahn, J.-H., Lee, J.-S., Yun, D.-W., Lee, M.-C., Eun, M.Y.
 and Kim, H.-I.
 Oryza sativa BAC OSJNB40028C22 genomic sequence
 Unpublished
 2 (bases 1 to 154345)
 Hahn, J.-H. and Kim, H.-I.
 Direct Submission
 Submitted (31-JAN-2002) Rice Genome Sequencing Project, National
 Institute of Agricultural Science and Technology (NIAST), RDA, 249
 Seodun-dong, Suwon 441-707, Korea (E-mail: jhahn@rda.go.kr,
 Tel: 82-31-290-0309, Fax: 82-31-290-0308)
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 2 contigs. Gaps between the contigs
 * are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * provided by the submitter.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.
 * 1 130145: contig of 130145 bp in length
 * 130146 130245: gap of unknown length
 * 130246 154345: contig of 24100 bp in length.
 Location/Qualifiers
 1. 154345
 /organism="Oryza sativa (japonica cultivar-group)"
 /variety="Nipponbare"
 /db_xref="taxon:39947"
 /chromosome="9"
 /clone="OSJNB40028C22"
 BASE COUNT 43614 a 31420 c 31651 g 45627 t 2033 others
 ORIGIN

Query Match 5.5%; Score 148; DB 2; Length 154345;
 Best Local Similarity 67.5%; Pred. No. 2.8e-31;
 Matches 208; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

QY 1405 CCTGAGGTCGTCATTTGTCAGCAAGACGATTTAAATATGTCCTAAACCTTACGCC 1464
 Db 144249 CCGGATGTTGTTATTTGTCCTGTAACCTGAGCTAGATATGCTGCAAAAGCATTTAGA 144190

QY 1465 GTTGCATTTGATGAGATCGTTTTTTTTCGTAACGACAGAGGAGGAAGATTTGCTCT 1524
 Db 144189 GACGCTTCAGGGTTGATATGCGTCTCTTCCAGATAGAGATGACTTTGTTCT 144130

QY 1525 TACAGTAAATTTCTCGGTAACCTTAAGAAAGATCGGGGGGTGTTCCAAATTAAGT 1584
 Db 144129 TACAGTAAATTTTGGCTACTTAAGGCTCAAAAGTCGGGAGGTTGTAAGGTTGAT 144070

QY 1585 GATGTACAACTTTATTTGTTGTCCTCATCAGATTTCTTAAGTATGATCTACCAAGTG 1644
 Db 144069 GGGGGGACACTTATTTTGGTCCACCTTCGATTTTTTTCGAAATGTTTTCGACAGTT 144010

OY	1464	CGTTCGCATGGATGAGATGCTTTTTTTTGCTACCAAGCAGGGAAGAATTGTGGC	1523
Dd	134307	CCAAATCATCTAAAGCATGGGTGGTTCCTTTCTTGTCTGGAAGATGCTGTAATCTGTT	134366
OY	1524	TTACACTGAATTTCTCCGTACCCTTAGCTCAACAAGATCGGGGGGTGGCCAAATTAGA	1583
Dd	134367	TTATGATAGTTTATGATCATTAATCTGGAGSAGAAAGCAGGGCAGCTTTCTTAATTTGA	134426
OY	1584	TGATGTGCACTTTAATCTTGTGGTGCCTCCATCAATGATTTCTTAAGTATGATGTAAGT	1643
Dd	134427	TGACCAACAACAGCTTTCTGGTGGTCCCTCCATCTGATTTCTCCGAGAAAGTACTTAAGT	134486
OY	1644	GACCCGTCAGAACGCCGTATATGTTGGTTGTTCCAAAGTA	1683
Dd	134487	TCCTGGGAACCTAGCATCTCGAGTTATTTCTAGCTTTA	134526
RESULT	10		
LOCUS	ATCHRIV34	196107 bp	DNA linear PLN 16-MAR-2000
DEFINITION	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 34.		
ACCESSION	AL161534		
VERSION	AL161534.2	GI:7267939	
KEYWORDS			
SOURCE			
ORGANISM	Arabidopsis thaliana. Arabidopsis thaliana.		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons: core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi 1 (bases 1 to 119463; 141613 to 196107) Peters,S.A., Van Steveren,M., Dirkse,W., Stiekema,W., Mewes,H.W., Lemcke,K. and Mayer,K.F.X. Unpublished 2 (bases 1 to 196107) EU Arabidopsis sequencing project. Direct Submission Submitted (10-MAR-2000) MIPS, D-82152 Martinsried, FRG, E-mail: Biochemie_Am_Klopferspitz_18a, D-82152 Martinsried, FRG, E-mail: lemcke@mips.biochem.mpg.de,mayer@mips.biochem.mpg.de Project Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK, E-mail: michael.bevan@bsrc.ac.uk Information on performance of analysis and a more detailed annotation of this entry and other sequences of chromosomes 3, 4 and 5 can be viewed at: http://www.mips.biochem.mpg.de/proj/thal/ this fragment has an overlap with ATCHRIV33 at the 5' end and an overlap with ATCHRIV35 at the 3' end. Location/Qualifiers 1..196107 /organism="Arabidopsis thaliana" /variety="Columbia" /db_xref="taxon:3702" /chromosome="4" 4530..23385 /note="Overlap to BAC T1P17; please refer to this entry for analysis and annotation of the gene fragment at the end of this BAC" 4533..23387 /note="Overlap to BAC T4C9; please refer to this entry for analysis and annotation" 5637..6595 /gene="At4g12380" complement(join(5637..5879,6044..6259,6572..6595)) /gene="At4g12380" complement(join(5637..5879,6044..6259,6572..6595)) /gene="At4g12380" /note="contains EST gb:T04352" /codon_start=1 /product="hypothetical protein" /protein_id="CAB78281.1" /db_xref="GI:7267940" /translation="MTYKKMLLHKRRKNKQMTETIIRNVPTPEIQEOCKRKEEKR		

LKSALEEGEGSDCFEFGSYDVSYNSGLINISMGVSYRFVIYTKKEIPWMNEDFLT
 KVTWGLGVQVALESYVGLMLCSSHQLVTRSPARHOVYISLPRHSARSPAVQLVVSRL
 "

exon complement(55637. .5879)
 /gene="AT4g12380"
 /number=1
 Intron complement(5880. .6043)
 /gene="AT4g12380"
 /number=1
 exon complement(6044. .6259)
 /gene="AT4g12380"
 /number=2
 Intron complement(6260. .6571)
 /gene="AT4g12380"
 /number=2
 exon complement(6572. .6595)
 /gene="AT4g12380"
 /number=3
 gene /gene="AT4g12390"
 8679. .9299
 CDS /gene="AT4g12390"
 /note="similarity to pectinesterase - Citrus sinensis,
 PID:g2098711
 Contains Prokaryotic membrane lipoprotein lipid attachment
 site AA34-44
 contains EST gb:AI992989.1, R90563, AI100571, T88173,
 H35998, AA067535"
 /codon_start=1
 /product="putative protein"
 /protein_id="CAB78283.1"
 /db_xref="GI:7267941"
 /translation="MEPKTLHCYCLILFLPLLCOSTIARPSSPNDSPSINFTVSSC
 RYMYQTLQYKCLAFADKTRRNNOIALQTLATLVAVOSTTYYGKLTARIRKRR
 EYLAVKDCVENLGLGELMLAOSKRELQVGSRGKRODELRLSNVEYWSAULTDT
 TCDLGFQKVDGVVKSARIRRVAVAVYSNALALVNRFAARRKS"
 8679. .9299
 /gene="AT4g12390"
 /number=1
 gene /gene="AT4g12400"
 complement(join(10992. .11105,11180. .11306,11384. .11693,
 11783. .12071,12529. .13365))
 /gene="AT4g12400"
 complement(join(10992. .11105,11180. .11306,11384. .11693,
 11783. .12071,12529. .13365))
 /gene="AT4g12400"
 /note="strong similarity to stress-induced protein still
 -Glycine max, PID:g872116
 contains EST gb:R64973, T20791, T22244, AA040981,
 AI100596, T46453, AI100738, T20868, AA395622"
 /codon_start=1
 /product="stress-induced protein still-like protein"
 /protein_id="CAB78283.1"
 /db_xref="GI:7267942"
 /translation="MAEEAKSGNAFSSGDYATITHTPEAINLSPTNHLIYSNRS
 SYSLHRYEALSDAKTIELKPDMSKYSRLGAFTLSKDEAVDSYKGLIDPS
 NEMLSGLADASBRVSSKSNPYDAPOGKEMKRLTRDPETRYLLEDDDDYKIMKEI
 QRPNNNLNLMKDRKRVKALGVLLNVAFGSSGSDTEKKEADEKKEPEPEPELVE
 EERQKRRERKALEKEGENVAYKKKDFGRAVEHYTKAMELDDDISLTVRAVYLE
 MKEYECCIECDKAVBERGRELSDFKMIALATPRKSALVMAKSCDFEALTFPOK
 ALTEHRNPDTLKLNDAEKYKLEOQEDPPTJAEERREKNGEFPEOKYPAIVKH
 SEATKRNPNRAVYASNAACYTKLGALPEGLKDAEKICELDSPSTKYSRKGALOFK
 KEYDKAMETVOEGIKHPKNOEFLDGYRRCVEQJNKASRGDLTPEELKERAKAMODP
 EVONILSDPYMRQVLVDFQENPKAOSHMKMPVMNIIQKLVSAGIVQR"
 complement(10992. .11105)
 /gene="AT4g12400"
 /number=1
 Intron complement(11106. .11179)
 /gene="AT4g12400"
 /number=1
 exon complement(11180. .11306)
 /gene="AT4g12400"
 /number=1

/number=2
 complement(11307. .11383)
 /gene="AT4g12400"
 /number=2
 exon complement(11384. .11693)
 /gene="AT4g12400"
 /number=3
 Intron complement(11694. .11782)
 /gene="AT4g12400"
 /number=3
 exon complement(11783. .12071)
 /gene="AT4g12400"
 /number=4
 Intron complement(12072. .12528)
 /gene="AT4g12400"
 /number=4
 exon complement(12529. .13365)
 /gene="AT4g12400"
 /number=5
 tRNA /number=5
 13550. .13623
 /product="tRNA-Asn"
 /note="codon recognized: GUU"
 15082. .15555
 /gene="AT4g12410"
 complement(15082. .15555)
 /gene="AT4g12410"
 complement(15082. .15555)
 /gene="AT4g12410"
 /note="similarity to auxin-induced protein 10A -Glycine
 max, PID:g255579
 contains EST gb:AI93046.1"
 /codon_start=1
 /product="putative protein"
 /protein_id="CAB78284.1"
 /db_xref="GI:7267943"
 /translation="MKRVRGFKIGHREYKIKWIPRNRCPTRTNPPYGINSLARCLS
 RGARLRCGSKKNGQNOIRLGDPKSKSNRYVPGHVLVHVGESDDPDTRVYVPVYF
 NHPLFGELEOAEVHGDPGRITPCRVSDFEKVQLRIAMPHCRKNSYKII"
 complement(15082. .15555)
 /gene="AT4g12410"
 /number=1
 gene /gene="AT4g12420"
 complement(join(22068. .22167,22248. .22447,22538. .22616,
 22697. .23062,23144. .23510,23593. .23731,23817. .24087,
 24195. .24304,24864. .24995))
 /gene="AT4g12420"
 complement(join(22068. .22167,22248. .22447,22538. .22616,
 22697. .23062,23144. .23510,23593. .23731,23817. .24087,
 24195. .24304,24864. .24995))
 /gene="AT4g12420"
 /note="EST GB:Z37701 marks 3' end
 strong similarity to pollen-specific protein precursor -
 common tobacco, PIR2:S22495
 contains EST gb:N38363, R30492"
 /codon_start=1
 /product="putative pollen-specific protein"
 /protein_id="CAB78285.1"
 /db_xref="GI:7267944"
 /translation="MDLEKILLFLEFVNISFCFADPYSFYNFVSYITASPLGVPQ
 VIALNGKPGPTIYNTNENLVVVRNKLDDGLLHNGIOQRVSMQDGYLGNCP
 PKNNMYTEPQVOKIGSFYEPSPILHQRASGGSPVNNRAIIPVDFSPDDITV
 TIGMYIRNTRKALKADDDGDLGPRGVLLNGGPRYRNTLVADGIDPETTYVHG
 KTYLRVSNVIGISLNRIGDHVLVLESGSTVOQNTYSLDIHVQOSTSLVTHD
 QMASDDYIVASARVNETIWRVTVGAILVYTSKSGAKAQDLPGGPDDDEDKFSN
 QARSIRNVSASGARPNPOGSEFKGSIIVTVVYLNNMPVTYIGSKRTTYLNGISFN
 PSTPIRLADKLKVDVYKLDPEKRPKPAVARSIIINGRGMEVYLONDRKMS
 YHMSGIAFFVVGMDYGEWTENSRGTYNMGDIARSTIOVYGAASAILIISDNGANN
 LRTEHDSWYIGQETVYRVVNDENKTERFHPDNLTCGALSRLQRPQKVSASAKS
 IGFYSLSNVVALVMMMLQH"
 complement(22068. .22167)
 /gene="AT4g12420"
 /number=1

25801	25000	gap of unknown length
25901	27016	contig of 1116 bp in length
27017	27116	gap of unknown length
27117	28154	contig of 1038 bp in length
28155	28254	gap of unknown length
28255	29648	contig of 1394 bp in length
29649	29748	gap of unknown length
29749	31439	contig of 1691 bp in length
31440	31539	gap of unknown length
31540	32736	contig of 1197 bp in length
32737	32836	gap of unknown length
32837	34705	contig of 1869 bp in length
34706	34805	gap of unknown length
34806	36569	contig of 1564 bp in length
36570	36469	gap of unknown length
36470	37820	contig of 1351 bp in length
37821	37920	gap of unknown length
37921	39065	contig of 1145 bp in length
39066	39166	gap of unknown length
39166	40330	contig of 1165 bp in length
40331	40430	gap of unknown length
40431	42100	contig of 1670 bp in length
42101	42200	gap of unknown length
42201	43481	contig of 1281 bp in length
43482	43581	gap of unknown length
43582	45285	contig of 1704 bp in length
45286	45385	gap of unknown length
45386	47189	contig of 1804 bp in length
47190	47289	gap of unknown length
47290	48669	contig of 1480 bp in length
48670	48869	gap of unknown length
48870	50612	contig of 1743 bp in length
50613	50712	gap of unknown length
50713	5217	contig of 2505 bp in length
53218	53318	gap of unknown length
53318	54410	contig of 1093 bp in length
54411	54510	gap of unknown length
54511	55559	contig of 1049 bp in length
55560	55660	gap of unknown length
55660	56993	contig of 1334 bp in length
56994	57093	gap of unknown length
57094	58794	contig of 1701 bp in length
58795	58894	gap of unknown length
58895	60398	contig of 1504 bp in length
60399	60498	gap of unknown length
60499	61727	contig of 1229 bp in length
61728	61827	gap of unknown length
61828	63020	contig of 1193 bp in length
63021	63120	gap of unknown length
63121	64540	contig of 1420 bp in length
64541	64640	gap of unknown length
64641	65992	contig of 1252 bp in length
65993	65992	gap of unknown length
65993	68071	contig of 2079 bp in length
68072	68171	gap of unknown length
68172	69597	contig of 1786 bp in length
69598	70057	gap of unknown length
70058	72094	contig of 2037 bp in length
72095	72194	gap of unknown length
72195	73893	contig of 1689 bp in length
73894	73984	gap of unknown length
73984	75384	contig of 1401 bp in length
75385	75484	gap of unknown length
75485	76673	contig of 1189 bp in length
76674	76773	gap of unknown length
76774	78925	contig of 2152 bp in length
78926	78926	gap of unknown length
78926	80869	contig of 1844 bp in length
80870	80969	gap of unknown length
80970	82475	contig of 1506 bp in length
82476	83575	gap of unknown length
83576	83646	contig of 1071 bp in length

	Query Match	1.7%;	Score 45.4;	DB 2;	Length 161477;
	Best Local Similarity	50.6%;	Pred. No. 0.2;		
Matches	137; Conservative	0;	Mismatches 131;	Indels 3;	Gaps 1;
QY	158 GAGGCTTGCGTTTATATACACAGCATGTGGAGAACAGTCGCACCCAAAGAGCCTC	217			
Dd	155197 GTGCCATTAGCCCTTCCTTAAGTTTCAGAACCTGACATGCGCACACAAGGCGCAAGSTGGCTA	155256			
QY	218 TTCAGAGAGCAAAATTTGATGTGAAGTCAAAATTAAGATGATACGACAGCACCGGGCAAAAC	277			
Dd	155257 TGTCTGGCCGGGTGATTTGGCACAACACCCTATAAAGTTAAGTGCTACG--GCAAGGCTTAACC	155313			
QY	278 CTTTAGAGAGTCTATGGGTGGGTGGATGGGCCCTATGCTCTCCAAGATGACTGGAGG	337			
Dd	155314 CTACACACCCGCTCTCGGGGTGGGTGTGCTTGGACCTAACACCTGCCTGGCGGCCCTGGCCA	155373			
QY	338 AAGAGTTACGCAAGTTTGGGAAAATCGAGAGATTTTAAAGTTTCTCAGAGAAGCCAGACAG	397			
Dd	155374 GAGAAATTCGATCGCTTTTGGAGACATTTGGACCATCGATCAGCTCAAAGAGACAGCTTTG	155433			
QY	398 CTTTCATTGATTATTTATGAGATGATGATGATGC	428			
Dd	155434 CCTACATCCAGTACGAGACCCTGGATGCTGACGC	155464			

RESULT 14			
AX083744			
LOCUS			
DEFINITION	AX083744	1141 bp	DNA
ACCESSION	Sequence 22 from Patent WO0111061.		
VERSION	AX083744		
	AX083744.1	GI:13185472	
			PAT 28-PEB-2001
			Linear

FEATURES	SOURCE
ORGANISM	synthetic construct
REFERENCE	artificial sequences.
AUTHORS	1 (bases 1 to 1141)
TITLE	Kunst,L. and Clemens,S.
JOURNAL	Regulation of embryonic transcription in plants
	Patent: WO 011061-A 22 15-FEB-2001;
	UNIVERSITY OF BRITISH COLUMBIA (CA)
	location/qualifiers

```

SOURCE      1. .1141
            /organism="synthetic construct"
            /db_xref="taxon:32630"
promoter    1. .1141

```

	BASE COUNT	ORIGIN
a	123	From local
c	32	Cuba
g	42	Greece
t	112	Turkey
others	832	Others

Query Match	1.7%	Score 45.2;	DB 6;	Length 1141;
Best Local Similarity	10.6%;	Pred. No. 0.21;		
Matches 82; Conservative	276;	Mismatches	413;	Indels 0;
				Gaps 0;

QY 379 CTGAGAGAGCGCAGACACACTTTCATTGATTTATATGAGATGATGATGCTTTACAGCT 438
Db 54 SKRSKWTVAARMYCYTRRWYNKSKRWKGGKKKKYBCAANTSBRAHARRKKDKMTAYBMT 113

QY 439 AAGAGCAGCAATGAAAGCCATGGGTGAGACCTTTTCGCGTGTGATTTTCGCGTCA 498
 Db 114 MTKNGKGTGWRHRYWRMBDVTVDHHVYTAAMNNAVTTCCMDXDDKRTFRFWMKNNNAAT 173

0y CAAGCGCCAAAAAAGAACAATGGGCTGGCTCTTAGCATAAACAGAAAATGGCAATGAAT 558
499 : : : : :
Db 174 GWDGDTKTHMNNNGCBVTWMAVYKTDDWSBKRNATGMBWMKNWSYVTYYWWVWDM 233

OY 559 CATAAACCGCAGATCCCTCACTCATATGAGAAGTCTTTAAAGAGAGATGTCCAGCCAGTAAGT 618
| : : : : : | : : : : : | : : : : : | : : : : :
Db 234 CCKKVRPRVTVTCGRMRTMYAMBEAHRRRYNNGTBAAMYRRTWTNNNNNNNAKMKCRK 293

QY 619 GGTCTGTGGATTGCGGTTCCCTCACTGCTACCAATGCAATGATGTGAGCCAAATCTCGAC 678

Db 294 YNCWNNRLADVNSTCTCTWKSCTTKYKRTSCAANNCRAGDANKDKHKWKWSAAMGVYNNNNNN 353

OY 679 AATGCGATGATACCTCTTGGTGAAGATCGAGAGGTAATAAGTTACCATCAAGCAATTTT 738
 Db 354 NMTYKKAHBAHBMWVWVHSAHMKWMAHNAHYSRKWTBYKRTVYNNNNGTMMRMAM 413
 OY 739 GCACCTTGAGGATTAGGAGCGCGGAGGAGACCTGCCAATGCAAGAGCGCTACAGGGG 798
 Db 414 YMKMDMBGTYNNNNNGGRTTYGWTAKKMTTYKKKAAHCKWPAHDKTCTHNTTW 473
 OY 799 AGTTATATCAATATCTAGATCAATAATATATGCTCAAGATGAGTGGCTCTGAG 858
 Db 474 KMTYNNCKYMSKSMNGSHRBAALVYTWMMWRHAYHANNNDYMMKACTWYKYBVC 533
 OY 859 CAAGACGATGATGTTTACTCTGCTGATGAACGCTCAAGACGATATGTTCAATAAT 918
 Db 534 SKMNNNYAMWTKSSWNTSRRYRKNTNNSWNRSDTSMGRANNYAABHGYKWNTRW 593
 OY 919 GATCCTTCATGTTCTCTCCATCTCTACTGTAATCCGAGGTCTATGAGGCCCTC 978
 Db 594 WBSHTTBHBRAGAHAHYMBMAYBAKCHMKAMWYAKKYAGAGSNNNNNNNNNNNN 653
 OY 979 AGAGTACGAATGAGCGCTTCAATATATGCTGAGATCAATGAGCTGTTGCTAGAGG 1038
 Db 654 NATCARDYAAASRWYAMANAAMKMYKBAANNAAYTHANNMGCNNATDERTMKNNN 713
 OY 1039 CCAACTGAGAGAGCGCATCTGCAAAATGGAATGCAATCTCCATCCCAACAGAGACT 1098
 Db 714 NNNAGTMMNNNNNNNAKNAKNAKNAKNAKNAKNAKNAKNAKNAKNAKNAKNAK 773
 OY 1099 GGAATCCCTCCATCTCTGCAAGAGTACGAGCGCCCTATGAGGTCAAC 1149
 Db 774 NGATYKTYTNNNTYRGVYNTATARDGMANNNNNNNNNNNNNNNNNNNNNNNN 824

RESULT 15
 AC094195
 LOCUS
 DEFINITION Rattus norvegicus clone CH230-2L6. *** SEQUENCING IN PROGRESS ***
 AC094195 152079 bp DNA linear HTG 09-JUL-2002
 AC094195 65 unordered pieces.
 AC094195.3 GI:21703371
 HTG: HTGS_PHASE1.
 KEYWORDS
 SOURCE Norway rat.
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 152079)
 Muzny,D.M., Adams,C., Adio-Oduola,B., All-oman,F.R., Allen,C.,
 Albrooks,S.L., Amaralunge,H.C., Are,J.R., Ayale,M., Banks,T.,
 Barbarta,J., Benton,J., Blinage,K., Blankenburg,K., Bonnin,D.,
 Bouck,J., Bowle,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P.,
 Bunay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
 Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
 Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
 Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
 Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
 Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
 Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
 Earhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
 Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
 Gabels,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
 Gorrell,J.H., Guevara,N., Gunaratne,P., Hale,S., Hamilton,K.,
 Harris,C., Harris,K., Hart,M., Haylak,P., Hawes,A., Hernandez,J.,
 Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
 Homsl,F., Howard,S., Huber,J., Hulik,S., Hume,J., Jackson,L.E.,
 Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
 Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
 Krtovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
 Li,J., Li,Z., Lichtarge,O., Lieu,X., Liu,J., Liu,W., Louised,H.,
 Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
 Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
 Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mel,G., Mettaker,M.,

REFERENCE
AUTHORS

TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
 Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
 Nguyen,N., Nickerson,E., Nwokenkwu,S., Ogun,M., Okunolu,G.,
 Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
 Peters,L., Plickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
 Rivers,M., Rojas,A., Rojokan,I., Rolfe,M., Ruiz,S., Savery,G.,
 Scherer,S., Scott,G., Shen,H., Shooshari,N., Sisson,I.,
 Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H.,
 Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
 Tansley,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
 Umanli,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wang,Q.,
 Wang,S., Ward-Moore,S., Warren,R., Washington,C., Wallington,S.,
 Williams,G., Williamson,A., Wleczek,R., Wooden,S., Worley,K.,
 Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorilla,S., Nelson,D.,
 Weinstein,G. and Glbbs,R.
 Direct Submission
 Unpublished
 2 (bases 1 to 152079)
 Worley,K.C.
 Direct Submission
 Submitted (17-SEP-2001) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 152079)
 Worley,K.C.
 Direct Submission
 Submitted (09-JUL-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Jul 8, 2002 this sequence version replaced gi:17940908.
 ----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GAEI
 Center clone name: CH230-2L6
 ----- Summary Statistics
 Sequencing vector: Plasmid
 Chemistry: Dye-terminator Big Dye: 100% of reads
 Assembly program: Phrap: version 0.990329
 Consensus quality: 99555 bases at least Q40
 Consensus quality: 103811 bases at least Q30
 Consensus quality: 107426 bases at least Q20

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 65 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 * 1
 * 1005 1004: contig of 1004 bp in length
 * 1105 1104: gap of unknown length
 * 2393 2392: contig of 1288 bp in length
 * 2492 2491: gap of unknown length
 * 2493 2492: gap of unknown length
 * 3557 3556: contig of 1065 bp in length
 * 3558 3557: gap of unknown length
 * 3658 3657: gap of unknown length
 * 4692 4691: contig of 1034 bp in length
 * 4791 4692: gap of unknown length
 * 6345 6344: contig of 1534 bp in length
 * 6346 6345: gap of unknown length
 * 7703 7702: contig of 1258 bp in length
 * 7704 7703: gap of unknown length
 * 7804 7803: gap of unknown length
 * 9117 9116: contig of 1314 bp in length
 * 9118 9117: gap of unknown length
 * 9218 9118: contig of 1251 bp in length
 * 10469 10468: gap of unknown length
 * 10568 10469: gap of unknown length
 * 11693 11692: contig of 1125 bp in length

```

* * * * * 75730 78546: contig of 2817 bp in length
* * * * * 78547 78646: gap of unknown length
* * * * * 78647 81125: contig of 2479 bp in length
* * * * * 81126 81225: gap of unknown length
* * * * * 81226 84905: contig of 3680 bp in length
* * * * * 84906 85005: gap of unknown length
* * * * * 85006 86613: contig of 1614 bp in length
* * * * * 86620 86719: gap of unknown length
* * * * * 86720 89463: contig of 2744 bp in length
* * * * * 89464 89563: gap of unknown length
* * * * * 89564 92460: contig of 2897 bp in length
* * * * * 92461 92560: gap of unknown length
* * * * * 92561 94856: contig of 2296 bp in length
* * * * * 94857 94956: gap of unknown length
* * * * * 94957 98221: contig of 3265 bp in length
* * * * * 98222 98321: gap of unknown length
* * * * * 98322 101646: contig of 3325 bp in length

Query Match 1.6% Score 44: DB 2: Length 152079;
Best Local Similarity 63.0%; Pred. No. 0.52; Mismatches 40; Indels 0; Gaps 0;
Matches 68: Conservative 0;

OY 154 TCACGAGGCTTTGGCTTTATATACACAGCATGTGAGGAGACAGTCGCGAGCCAAAGAG 213
Db 16271 TCAAGAGGATTTGCCCTTTGATATTTTGAAGAAATGATGATGATGCCAAGGAAGCTAAGAA
OY 214 GCTCTTCAGGAGCAAAATTGATGATGAGCAAGCAAAATTAAGATCGAATAC 261
Db 16331 CGTGCCAAATGGAATAGAGCTTGATGAGCGTGAATTAAGGCGCATTTTC 16378

Search completed: June 19, 2003, 03:22:38
Job time : 6977 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 19, 2003, 00:45:48 ; Search time 3748 Seconds
(without alignments)
11692.899 Million cell updates/sec

Title: US-09-920-705-2
Perfect score: 2706
Sequence: 1 atggcgtactatgaagcc.....cgctgacagggccttga 2706

Scoring table: IDENTITY-NUC
Gapop 10.0, Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST: *
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vit:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	403.6	14.9	701 10	BE037904 AA06B09 A
2	403	14.9	767 17	BH552322 BOGH093TF
3	350.2	12.9	635 9	AU237346 AU237346
4	334.6	12.4	724 17	BH552335 BOGH093TR
5	330.8	12.2	602 12	BG543956 E1706 Ch1
6	323.4	12.0	416 17	BH851246 SALK_0727

7	303	11.2	582	10	AV829561
8	194	7.2	561	12	BF266387
9	177.8	6.6	655	10	AM96401
10	176	6.5	630	12	BG447636
11	170.8	6.3	647	14	BO696544
12	170.8	6.3	662	12	BF641487
13	170.2	6.3	685	13	B1954416
14	165.6	6.1	617	12	BG649121
15	161	5.9	491	10	AV548323
16	154.2	5.7	381	14	BQ697581
17	145.4	5.4	606	10	BE594381
18	139.8	5.2	493	12	BF639297
19	133.4	4.9	472	14	N96947
20	129.8	4.8	691	9	AL505369
21	123.2	4.6	673	12	BG600897
22	121.8	4.5	429	10	AV406694
23	118.6	4.4	751	12	BG591991
24	118.4	4.4	317	14	C24205
25	112.8	4.2	430	12	BF585815
26	112	4.1	475	9	AU228388
27	108	4.0	526	10	AM033033
28	106.6	3.9	423	10	AV407805
29	100.6	3.7	589	12	BG649379
30	96.4	3.6	644	12	BF278330
31	95.4	3.5	670	14	BQ401768
32	95	3.5	586	10	BE499504
33	91.2	3.4	820	17	BH715029
34	85.4	3.2	539	10	AM596129
35	84.6	3.1	457	9	AL378093
36	84.6	3.1	473	9	AL378092
37	80.8	3.0	439	14	BO489608
38	76.6	2.8	480	10	AV834854
39	73.2	2.7	334	13	B1751900
40	72	2.7	577	12	BF003832
41	71.8	2.7	493	10	AV529289
42	71.8	2.7	528	9	AU237648
43	69.4	2.6	430	17	BH251458
44	65.4	2.4	475	9	AU174330
45	62.8	2.3	769	17	BH442944

ALIGNMENTS

RESULT 1
LOCUS BE037904 701 bp mRNA linear EST 07-JUN-2000
DEFINITION AA Arabidopsis thaliana cDNA 5', mRNA sequence.
ACCESSION BE037904
VERSION BE037904.1 GI:8332920
KEYWORDS EST.

SOURCE
ORGANISM Arabidopsis thaliana
thale cress.

REFERENCE
AUTHORS Bohnerl,H.J., Borchert,C., Brazille,S., Brooks,J., Eaton,M., Ferreira, H., Kawasaki,S., McColough,A., Michalowski,C.B., Palacio,C., Scara,G., Wheeler,M. and Zepeda,G.R.
Functional Genomics of Plant Stress Tolerance
Unpublished (2000)
Contact: Michalowski,C.B.
University of Arizona
Bio Sciences West room 513, Tucson, AZ 85721, USA
Tel: 520-621-7982
Fax: 520-621-1697
Email: chmdu.arizona.edu

TITLE
JOURNAL
COMMENT
Best Blastx match: 'gb|AA064314.1 (AC002335) hypothetical protein [Arabidopsis thal. . 165 3e-40'. An open reading frame exists.
Insert Length: 2 Std Error: 0.00.
Location/Qualifiers
1. .701

FEATURES
SOURCE

Email: catowner1@rtr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TF

Db 40 AACATGCCCTCAGATGATC 22

[illegible][illegible]

QY		903	AGATATGTCATATATGATTCCTCATATGTATATCTTCCATTCTACTGTAATCTCCGG	962
Dd		529	AGGATGTTTCATCACTGAATCCTTCTAT	554
QY		963	GTCATATGAGGCCCTCAGAGGTACGAATGAGCGTTTCATATATATGTCGAGAATAATGA	1022Z
Dd		555	-----TGAAAGGCATTCACAGAGGTAGCAATGAGACCCTTATATATGTTTCAATGA	609
QY		1023	CGTTGTGTAGTAGAGCCCAATGAGAGAGGCCATCTGCAATATGAAATGCAATATCTCCC	1082Z
Dd		610	CGTTGTGTATGATGAGGCCAATAATGAGGAGCGCCGTTCCAATATGAGGATGGGATATCTCC	669
QY		1083	ATCTCCAACAGAGACCTTGGAATCTCTCCATCTCTCCACACAAGATGACAGGCGCCTT	1137
Dd		670	ATCTCCACAGGACATATGAGATCCCTCCGCTCTCTCCACACAAGATGAGAACCTT	724
RESULT 5				
BGS43956				
DEFINITION		BGS43956	602 bp	mRNA linear EST 01-MAY-2002
LOCUS		E1706	Chinese cabbage etiolated seedling library Brassica rapa	
ACCESSION		BGS43956		
VERSION		BGS43956.1	GI:20374936	
KEYWORDS		EST.		
SOURCE		Brassica rapa subsp. pekinensis.		
ORGANISM		Brassica rapa subsp. pekinensis		
REFERENCE		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
AUTHORS		Ryu,S.H., Kang,U.S., Kang,C.-h., Kim,C.Y., Choi,Y.J., Lee,S.-H.,		
TITLE		Bank,J.D., Lee,S.Y., Cho,M.J. and Lim,C.O.		
JOURNAL		Expressed Sequence Tags of Chinese Cabbage Etiolated Seedling cDNA		
COMMENT		Unpublished (2001)		
		Contact: Lim, C.O.		
		Plant Molecular Biology & Biotechnology Research Centre		
		Gyeongsang National University		
		#900 Gazwa-dong, Jinju 660-701, Korea		
		Tel.: 82 35 751 6255		
		Fax: 82 35 759 9363		
		Email: colim@hongae.gsnu.ac.kr		
		Seq primer: T7.		
FEATURES				
SOURCE		Location/Qualifiers		
		1..602		
		/organism="Brassica rapa subsp. pekinensis"		
		/cultivar="Jangwon"		
		/db_xref="taxon:51351"		
		/clone="E1706"		
		/clone_lib="Chinese cabbage etiolated seedling library"		
		/tissue_type="Etiolated seedling"		
		/lab_host="XL-1 Blue"		
		/note="Vector: pSPORT 1; Site_1: Sal I; Site_2: Not I"		
BASE COUNT		147 a 144 c 166 g 145 t		
ORIGIN				
Query Match		12.2%; Score 330.8; DB 12;		Length 602;
Best Local Similarity		75.6%; Pred. No. 1,3e-91;		
Matches		440; Conservative 0; Mismatches 132; Indels 10; gaps 2;		
QY		67	GGTAGCCTAAGCCCGGAGACGACAGAGTCAATCGAGTGTGTTGGAAATACGCG	126
Dd		13	GGCAGCTCGACCCCGGACTACTACCGATCGCATCGCATCGCATCTGGGAAGTTCCGC	72
QY		127	GATATGATATGAATACAGGCTGATATCTTACACAGAGCTTTCGTTTATATCTACACAT	186
Dd		73	GATATGACAGGCTCACACCGCTATCTTCCCGTTGCTTCGTTTATCTTACCTCAGCGC	132
QY		187	GTCGAGACACAGTCGACGACCAAGAAGGCTTTCACAGACCAATTTGAATGAGATCAA	246
Dd		133	GTTGAGGAACCGGAGCGCGGAAGAGGCTCTTCAGGTCCTCAATCTAAACGGAGTCCG	192
QY		247	ATTATAGATGCAATATACGACGACCGCAAAACCTTGTAAGAGCTATAGGCTGGGTGAATC	306

Db	193	ATTAAAAATCGATTTCGACAGACCGGCAAAACCTTTGTAAGACTCTTTGGTTGGTGAATC	252
QY	307	GGCCCTAATGTCCTCCAGAGATGACCTGGAGGAAAGATTACAGACAGTTTGGGAAATTCAG	366
Db	253	AGCTGAGATGTCGCGAAGAGATGTACTGACGAGCAGATTACAGAGATTGGGGAATTCGAA	312
QY	367	GATTTAGGTTTTCACAGAAACGACAGACGCTTCATGTTATTAAGANGATGAT	426
Db	313	AGTTTCAGGTTTCTTAAGAGATCGGACAGACGCTTTTATGACTTTTTCGACGTGGACGAT	372
QY	427	GCTTACAGGCTAAGAGAGATGAATGAAGAAAGCTTATGGGTGTACTTTTGGCGTTGAT	486
Db	373	GCTGACAGCGAAAGACCATATAGCGGAGAGCATAGTGTGACTTTTTCGAGATTGAT	432
QY	487	TTTTCGCGGTACAGACGCCCAAAAAAAGAACATGGGCTGGCTTTCAGATTAACGAAAT	546
Db	433	TTTCTTCGCTCACAAGGGCCAAAGAAAGAACACAGCGTTGGCTCCGAAAGATGCGAGGAT	492
QY	547	GGCAATATGATCATTAACCGCAGTATGCTCATCATATGAAGACTTAAAGAGATGTC	606
Db	493	GGCAGTTGATGTGCTAGCACACACA-----GTAATCTACTCATCTAGAGATGTC	543
QY	607	CAGCCAGTAAGGTTCTGTGTGATTTGGTTGGTTCCCTCTACTGCT	648
Db	544	CA-ACAAGCAAGTTCTGTGTGATTTGGTGGTACCTCTCTCTGT	584
RESULT 6			
LOCUS	BH851246		
DEFINITION	SAIK_072730.56.00.x Arabidopsis thaliana TDNA insertion lines		
ACCESSION	BH851246		
VERSION	BH851246.1		
KEYWORDS	GI:21422117		
ORGANISM	thale cress.		
REFERENCE	Arabidopsis thaliana		
AUTHORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons: core eudicots; Rosidae; eucotids II; Brassicales; Brassicaceae; Arabidopsis. 1 (bases 1 to 416)		
TITLE	Alonso,J.M., Leisner,T.J., Barajas,P., Chen,H., Cheuk,R., Gadriab		
JOURNAL	,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P.		
COMMENT	Zimmerman,J. and Ecker,J.R.		
SOURCE	A sequence-indexed library of insertion mutations in the Arabidopsis genome		
	Unpublished (2001)		
	Contact: Joseph R. Ecker		
	Salk Institute Genomic Analysis Laboratory (SIGNAL)		
	The Salk Institute for Biological Studies		
	10010 N. Torrey Pines Road, La Jolla, CA 92037, USA		
	Tel: 858 453 4100 x1752		
	Fax: 858 558 6379		
	Email: ecker@esalk.edu		
	This is single pass sequence recovered from the left border of TDNA. This sequence lies within an annotated exon of At2g343410.		
	Class: TDNA tagged.		
FEATURES	location/Qualifiers		
source	1..416		
	/organism="Arabidopsis thaliana"		
	/strain="Columbia 0"		
	/db_xref="taxon:3702"		
	/clone="SAIK_072730.56.00.x"		
	/note="PCR was performed on Arabidopsis thaliana lines each of which contains one or more TDNA insertion elements. The resultant fragment for each line was directly sequenced to determine the genomic sequence at the site of insertion. Details of the protocols used can be found at http://signal.salk.edu/tdna_protocols.html ."		
BASE COUNT	136 a 109 c 78 g 93 t		


```

/adb_xref="taxon:4513"
/clone="HV_C850015A03f"
/clone_lib="Hordeum vulgare seedling green leaf EST
library HVCNDN0004 (Blumeria challenged)"
/tissue_type="seedling green leaf"
/lab_host="TJC121"

```

/note=Vecori: lambdaZAP; Site.1: SCOR1; Site.2: Xho1; C.I. 16155 (Mal3) plants were greenhouse grown in the R Wise lab at Iowa State University, Ames, IA; 7 day old greenhouse seedlings were challenged with isolate A27 (AYRMA1a13) of *Blumeria graminis* f. sp. hordei, and leaves were harvested 20 and 24 hr post-inoculation and snap frozen; uninoculated leaves were harvested 20 hr post-inoculation (Wei, Wise). In the TV Close lab at the University of California, Riverside, total RNA was prepared from each sample pool, equal quantities of all three RNA pools were combined, poly(A) RNA was purified from the mixture, one cDNA library was made, and 1 million pfu were in vivo excised to give pBluescript SK(-) cDNA phageids (Choi, Close). Phageids were plated and picked at the Clemson University Genomics Institute (CUGI) (Begum, Palmer, Frisch, Atkins and Wing). Plasmid DNA preparations, DNA sequencing and sequence analysis were performed at CUGI (Wing, Yu, Frisch, Henry, Slimmons, Oates, Rambo, Main). The sequence has been trimmed to remove vector sequence and contains a minimum of 100 bases of phred value 20 or above. For more details on library preparation and sequence analysis see

Genetically and Physically anchored EST resources for barley genomics. Barley Genetics Newsletter 31:29-30. (<http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html>)" see Close TJ, Wing R, Kleinofhs A, Wise R (2001)

BASE COUNT	130 a	117 c	149 g	165 t
ORIGIN				

Query Match	7.2%;	Score 194;	DB 12;	Length 561;
Best Local Similarity	68.3%;	Pred. No. 9.7e-49;		
Matches 269; Conservative	0;	Mismatches 125;	Indels 0;	Gaps 0;

QY 1311 TGAATCATATGGAGAGCAATGATTGCCAAGGTTGGAACCTCCCGTGTGTGCTCGTTG 1370

0V TCAGGACCTATTCGGCAACCACCATCACAATTCCAAACTGCAGCGTGTGCCAAGAAGTTTGGAATAAGTA
26 GAGTGGGGCACATCTGGGCGTAGCCAGATATTCCTCAAAGTAGAACATCTATGGTTTTGTCGTGTCTGGGTG 14 /

Dp 148 CTTATCTATAAGAGAGGGCGCTTGAGATACCTTTACGCGACATTTTAATAATCTCTGCTAG 207

QY 1431 AACTGATTTCATATGCTGCTAAACATTACGCCGTTGCCATTGGATGTGAGATCGTTTT 1490

[illegible]

Dp 268 CTTCTACAGACAGTGAAGATGATTTGTGCTCTTACACTGAATTTCTGGCGTACTTGGG 327

QY 1551 CTCAAAAGATCGGGCGGTGTCCAAATTAGATGATGTACACTTTATTCTTGTGCC 1610

[illegible]

Db
388 GCCATCGGATTTCTGACGAATGTTCTGCAGTTGATGCTCCAGAGCGCCCTTTATGCGCT 447

QY 1671 TCTTCTCAAGTTACCCCCGCCAGCCGTTCTGT 1704

448 GGTACTGCACATTCACAGATGTCCATGTCTGCT 481

RESULT 9
AW696401

[illegible]

DEFINITION	NP103EB7/ST1P1054 Developing stem <i>Medicago truncatula</i> cDNA clone
ACCESSION	NP103EB7/ST 5', mRNA sequence.
VERSION	AW696401
KEYWORDS	AW696401.1 GI:7571086
SOURCE	EST.
	barrel medic.

ORGANISM	REFERENCE
Medicago truncatula	1 (bases 1 to 655)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabcaceae; Papilionoideae; Trifoliales; Medicago.	1 (bases 1 to 655)

AUTHORS
 He, A.-Z., Snadole, G., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell,
 C.-J., Flores, H.R., Inman, J.T., Weller, J.W., May, G.D. and Dixon,
 R.A.
 TITLE
 Expressed Sequence Tags from the Samuel Roberts Noble Foundation
 Medicago truncatula stem library
 JOURNAL
 Unpublished (2000)
 COMMENT
 Contact: Dixon RA
 Plant Biol. Division

Plant Biology Division
The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73402, USA
Tel.: 580 221 7302
Fax: 580 221 7380
Email: radix@noble.org
Insert length: 655 Std Error: 0.00
Plate: 103 row: E column: 07
Seq primer: TCACACAGGAACGCTTATGAC.

FEATURES	SOURCE	Location/Qualifiers
		1. .655
		/organism="Medicago truncatula"
		/db_xref="taxon:3880"
		/clone="NF103E075P"
		/clone_lib="Developing stem"
		/tissue_type="stem"
		/dev_stage="Pooled developmental"
		/note="Vector: Lambda Zap; Contains a mixture of internodal stem segments"
BASE COUNT		182 a 149 c 137 g 187 t
ORIGIN		

Query Match	6.6%	Score 177.8;	DB 10;	Length 655;
Best Local Similarity	62.6%	Pred. No. 1.3e-43;		
Matches 311; Conservative	0;	Mismatches 182;	Indels 4;	Gaps 2;

QY 29 CCGATGATTCCGGTTTCCAGTCACAACAACTTTGGTGCGTAGCCTAACGCCCGGAGACGA 88
| | | | |
Db 159 CGGAGGATTTCTTGCAGCTCTTCCAACAACCTATATGTTCGCAATCTAATGCCGAGACTTA 218
| | | | |

89 CAGAGTCAGATCTGACCCAGTTGTTGGAGATACGGCGATATTGATTCGATCAGCGGTCT 148

DD 219 CAAATTCGACGATGGATCTCTCGTCAGTACGGGTCTCTCGATAGCGGTCAATCTT 278
 QY 149 ATTCTTACGAGGCTTTCGTTATATATCTACAGACATGTGGAGAAACAGTCTGACGCCA 208

Dd 279 ATTCCGGTCGCAATTACGCTTTCGTTCCTTCAAAAGCATTTGATGATGCCAAAGCTGCTA 338

Oy 209 AAGAGGCTTCTCAAGAGCAAAATTTGATGATGAGTCAAATTAAGATCGATACGCACGAC 268

Db 339 AAAACGCTCTTAGGGATTCAAATTTTCGTGGCAATTCCTTGAACAATTTGAGTTTGCTAAC 398

[illegible]

Db

459 ACTTGGAGCTGACTTCGGTAAATTTGTTAAAGTTGAGGATTCCAAGTTCTTCAGAGACA 518

389 GCAAGACAGCCTTCATTCGATTTTATGATGATGATGATGCTTACAGC---TAAGGCA 445
 519 GAAATCTGCATGTGTCGAATTTTTCATTTTATGATGATGCTTATTCAGCAATGAAAAATTA 578

QY	1235	CTCCAAATGGGTGTCGATGAGAGTCAATTTGGTGCAGGTTCACTGTCCTAGACCTATACC	1294
	1111		1111
Db	259	ATCCTTATGACAGGAGGAGGTATCCTTGCTTCACGACACGAGGATTAACAGGTGAGTAC	318
QY	1295	GTGGCCCCCTGATTTCTGATTCACATATATGAGAGAAATGATTTGCCAAGGTTGAACTCCCG	1354
	1111		1111
Db	319	ATGCTGATCC--AACCTGATCATATATATGCGCGGACATTAATGCAAAAGGAGAACTCTCG	375
QY	1355	TCTGTTTGTCGTCGTTGTTGTAACCTATATGGGAAAGGAGGATTTGAAACATTAACCTGCTAGGTCG	1414
	1111		1111
Db	376	TTTTCGTGCTGATGATGATCCCTTATAGGAAAGGATTCGGACAGACACTTCTTAAGTTG	435
QY	1415	TCAATTTGTCAGCAAGAACATGATTTGAATATGTCGTAAACATTTACCGCTTGCCATTG	1474
	1111		1111
Db	436	TGATATCTCTGACGTAGACGGGATTTGATACACTCGCAGCACACTATGCTGATCAATTG	495
QY	1475	GATGTGATGATCGTTTTTTCTTCCATACCACAGGAGAAAGATTTTGCTCTTACACTCAAT	1534
	1111		1111
Db	496	ACTTTGAGATGTTTCTTCTGCTATAGTGAATAATGATTTGGTTGATCACTGAAT	555
QY	1535	TTTCCCGGTAACCTTAGCT--CAAAAGATTCGGGCGGTGTCGCAATTAAGATGATGATACA	1593
	1111		1111
Db	556	TTCTACACGCTACCTTTGGTGGCAAAATTCGTCTGTGTTGCCAATTTTGGAGATACA-A	614
QY	1594	ACTTTATTTCTTGCTGTC	1609
	1111		1111
Db	615	ACTTTATTTCTTGCTGTC	630
RESULT 11			
B0696544		647 bp	mRNA linear EST 15-JUL-2002
LOCUS	NPXP_042.E01.F	NPXP (Nsf Xylem Planlings wood Vertical)	Pinus taeda
DEFINITION	CDNA clone NPXP_042.E01.5', mRNA sequence.		
ACCESSION	B0696544		
VERSION	B0696544.1	GI:21821860	
KEYWORDS	EST.		
SOURCE	Pinus taeda		
ORGANISM	loblolly pine.		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Coniferales; Pinaceae; Pinus; Pinus.		
AUTHORS	Sederoff, R.		
TITLE	Molecular Basis of Wood Formation in the Pine Megagenome		
JOURNAL	Unpublished (2000)		
COMMENT	Contact: Johnson, Arthur North Carolina State University Tel: 919 515 7800 Fax: 919 515 7801 Email: ajohnson@ncsu.edu		
FEATURES	Seq primer: T3:		
Source	Location/Qualifiers		
	1. 647		
	/organism="Pinus taeda"		
	/strain="Coastal plain loblolly pine from North Carolina"		
	/db_xref="taxon:3352"		
	/clone="NPXP_042.E01"		
	/clone_1bp="NPXP (Nsf Xylem Planlings wood Vertical)"		
	/tissue_type="Xylem"		
	/cell_type="Planlings (secondary)"		
	/dev_stage="Transitional"		
	/lab_host="X11-Blue"		
	/note="Vector: Bluescript SK; Site_1: Eco RI; Site_2: XhoI; The library is from early (spring) secondary wood, taken from a ten year old tree in the transitional phase. The tree is a kind gift of the Westvaco Corporation. Secondary xylem was harvested from the tree by peeling back the bark and primary xylem and then removing the underlying tissue with a block plane. NOTE: The sequences contain a 'CDNA adapter' between the EcoRI site and the start of the EST. The adapter sequence is 'AATTCGACGACGAG'."		
BASE COUNT	177 a	137 c	136 g
ORIGIN			172 t
			25 others

[illegible]


```

Db      182  AAATT---TGAGATACAACTTATCTGTGGTGCCTTCAGATTCCGACGAGTGT 238
QY      1636  CTCCAATGACCCCTCAAGAACGGCTATATGTTTCTCTCAAGTTACCCCGCCAGCC 1695
Db      239  CTGAAGTCACCCGACCGAAGCTATATGTTGTCCTCAAGTTGCACCGACATCAA 298
QY      1696  GTTCCCTGTACAGCATCATACAGAACAAATCTCAGTCACATCCTTCATATATGAT 1755
Db      299  AGTGTGTGTGCCCGTCACCATCATCATTTGCTGTACCGTCAATCATATATATGAC 358
QY      1756  CAGCCCGGAGTTCACCTGC 1775
Db      359  CAGATGCCTCTTCACAGGC 378

```

RESULT 15

AV548323/c

LOCUS AV548323 491 bp mRNA linear EST 06-SEP-2000
 DEFINITION AV548323 Arabidopsis thaliana roots Columbia Arabidopsis thaliana
 CDNA clone RZL51908F 3', mRNA sequence.

ACCESSION AV548323

AV548323.1 GI:8719736

VERSION EST

AV548323.1 GI:8719736

KEYWORDS

EST

SOURCE

thale cress.

ORGANISM

Arabidopsis thaliana

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

COMMENT

Contact: Erika Asamizu

The First Laboratory for Plant Gene Research

Kazusa DNA Research Institute

Yana 1532-3, Kisarazu, Chiba 292-0812, Japan

Email: asamizuekazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

FEATURES

source

1.491

/organism="Arabidopsis thaliana"

/strain="Columbia"

/db_xref="taxon:3702"

/clone="RZL51908F"

/clone_lib="Arabidopsis thaliana roots Columbia"

/tissue_type="roots"

/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2: XhoI"

BASE COUNT

150 a 92 c 96 g 153 t

ORIGIN

Query Match

Best Local Similarity

100.0%; Score 161; DB 10; Length 491;

Matches 161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2546 AGCCTATGATCTCTGGGATGTGTCAGGATACACAGATGGGAGGTCGATTAAGATCAGA 2605

Db 491 AGCCTATGATCTCTGGGATGTGTCAGGATACACAGATGGGAGGTCGATTAAGATCAGA 432

2606 GATACAGTCACACTACATTTGTCAGCAAAACCTTCTCTCCAGATACAGCAAAACAGC 2665

Db 431 GATACAGTCACACTACATTTGTCAGCAAAACCTTCTCTCCAGATACAGCAAAACAGC 372

2666 AGCAACAGTCTAGTACTCCGGCTGCAGAGGCGCTTGA 2706

QY 371 AGCAACAGTCTAGTACTCCGGCTGCAGAGGCGCTTGA 331

371 AGCAACAGTCTAGTACTCCGGCTGCAGAGGCGCTTGA 331

Db 371 AGCAACAGTCTAGTACTCCGGCTGCAGAGGCGCTTGA 331

371 AGCAACAGTCTAGTACTCCGGCTGCAGAGGCGCTTGA 331

QY 371 AGCAACAGTCTAGTACTCCGGCTGCAGAGGCGCTTGA 331

371 AGCAACAGTCTAGTACTCCGGCTGCAGAGGCGCTTGA 331

Db 371 AGCAACAGTCTAGTACTCCGGCTGCAGAGGCGCTTGA 331

371 AGCAACAGTCTAGTACTCCGGCTGCAGAGGCGCTTGA 331

QY 371 AGCAACAGTCTAGTACTCCGGCTGCAGAGGCGCTTGA 331

371 AGCAACAGTCTAGTACTCCGGCTGCAGAGGCGCTTGA 331

Db 371 AGCAACAGTCTAGTACTCCGGCTGCAGAGGCGCTTGA 331

371 AGCAACAGTCTAGTACTCCGGCTGCAGAGGCGCTTGA 331

QY 371 AGCAACAGTCTAGTACTCCGGCTGCAGAGGCGCTTGA 331

371 AGCAACAGTCTAGTACTCCGGCTGCAGAGGCGCTTGA 331

Db 371 AGCAACAGTCTAGTACTCCGGCTGCAGAGGCGCTTGA 331

371 AGCAACAGTCTAGTACTCCGGCTGCAGAGGCGCTTGA 331

QY 371 AGCAACAGTCTAGTACTCCGGCTGCAGAGGCGCTTGA 331

371 AGCAACAGTCTAGTACTCCGGCTGCAGAGGCGCTTGA 331

Search completed: June 19, 2003, 04:25:36
 Job time: 3767 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 19, 2003, 03:22:45 ; Search time 914 Seconds
(without alignments)
4344.484 Million cell updates/sec

Title: US-09-920-705-2

Perfect score: 2706

Sequence: 1 atggcgctatctatgaagcc.....cgcgtgacagggccttga 2706

Scoring table: IDENTITY_NUC

Searched: Gap0 10.0 , Gapext 1.0

Total number of hits satisfying chosen parameters: 2085038

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCr_NEM_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEM_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEM_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCrUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEM_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEM_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEM_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEM_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. NO. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	2706	100.0	2706	9	US-09-920-705-2
2	1896.8	70.1	4593	9	US-09-920-705-1
3	797.8	29.5	801	9	US-09-920-705-4
4	271.4	10.0	349	9	US-09-920-705-5
5	271.4	10.0	3715	9	US-09-920-705-6
6	113.2	4.2	438	10	US-09-770-444-972
7	42.8	1.6	256	10	US-09-815-343-1381
8	41.6	1.5	2142	10	US-09-917-800A-1567
9	39.6	1.5	653	9	US-10-123-155-438
10	38.6	1.4	1344	12	US-10-062-254-287
11	38.2	1.4	1049	9	US-10-123-155-358
12	37.2	1.4	424	10	US-09-878-574-4075
13	37.2	1.4	1537	12	US-10-062-254-275
14	37	1.4	1846	9	US-10-197-666A-77
15	37	1.4	1964	9	US-10-197-666A-79
16	37	1.4	2037	9	US-10-071-766-104
17	36.4	1.3	396	9	US-09-803-719-708
18	36.4	1.3	406	9	US-09-803-719-783
19	36.4	1.3	407	9	US-09-803-719-781

20	36.4	1.3	466	9	US-09-918-995-11693	Sequence 11693, A
21	36	1.3	1024	9	US-10-123-155-198	Sequence 198, App
22	36	1.3	2000	9	US-09-938-842A-2746	Sequence 2746, Ap
23	35.8	1.3	354	9	US-10-102-524-1033	Sequence 1033, Ap
24	35.2	1.3	637	10	US-09-778-320-189	Sequence 189, App
25	35.2	1.3	637	10	US-09-910-689-189	Sequence 189, App
26	35.2	1.3	637	12	US-10-010-742-189	Sequence 189, App
27	35.2	1.3	1583	12	US-10-062-254-273	Sequence 273, App
28	35	1.3	1799	12	US-10-062-254-247	Sequence 247, App
29	35	1.3	1806	12	US-10-062-254-249	Sequence 249, App
30	34.8	1.3	487	9	US-10-123-155-528	Sequence 528, App
31	34.8	1.3	1184	9	US-10-123-155-412	Sequence 412, App
32	34.8	1.3	1728	9	US-09-794-589-3	Sequence 3, Appl
33	34.4	1.3	734	9	US-10-184-644-458	Sequence 458, App
34	34.4	1.3	734	9	US-10-184-644-458	Sequence 458, App
35	34.4	1.3	2552	10	US-09-799-777-141	Sequence 141, App
36	33.8	1.2	407	9	US-09-918-995-36239	Sequence 36239, A
37	33.8	1.2	600	10	US-09-917-800A-857	Sequence 857, App
38	33.6	1.2	867	9	US-10-123-155-20	Sequence 20, Appl
39	33.6	1.2	1160	9	US-10-123-155-234	Sequence 234, App
40	33.6	1.2	1757	12	US-10-062-254-269	Sequence 269, App
41	33.4	1.2	421	10	US-09-864-761-265	Sequence 265, App
42	33.4	1.2	470	10	US-09-864-761-867	Sequence 867, App
43	33.4	1.2	479	10	US-09-864-761-794	Sequence 794, App
44	33.4	1.2	581	10	US-09-864-761-9417	Sequence 9417, Ap
45	33.4	1.2	2116	12	US-10-002-600-81	Sequence 81, Appl

ALIGNMENTS

RESULT 1
US-09-920-705-2
Sequence 2, Application US/09920705
Publication No. US20030079252A1
GENERAL INFORMATION:
APPLICANT: Amasino, Richard M.
APPLICANT: Schomburg, Fritz M.
APPLICANT: Michaels, Scott D.
APPLICANT: Patton, David
TITLE OF INVENTION: Floral Induction Gene
FILE REFERENCE: 960296.97214
CURRENT APPLICATION NUMBER: US/09/920,705
CURRENT FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 2706
TYPE: DNA
ORGANISM: Arabidopsis thaliana
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(2706)
NAME/KEY: misc_binding
LOCATION: (46)..(279)
OTHER INFORMATION: RNA Binding Region
NAME/KEY: misc_binding
LOCATION: (283)..(522)
OTHER INFORMATION: RNA Binding Region
NAME/KEY: misc_binding
LOCATION: (610)..(852)
OTHER INFORMATION: RNA Binding Region
US-09-920-705-2
Query Match 100.0%; Score 2706; DB 9; Length 2706;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2706; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCGTATCTATTAAGCCATTCAGACCCATTCCTCCGTTTCCATCAACAATCTT 60
DB 1 ATGGCGTATCTATTAAGCCATTCAGACCCGATGTTCCGTTTCCATCAACAATCTT 60
QY 61 TGGCGGTAGCGCTTAACCCGAGACAGACAGATCAGATCTGACCGACTTGTGGAGCA 120

Db 61 TGGGCGGTAGCCTAACCCCGAGACGACAGTCAGATCGACCGAGTCTTTGGAGA 120
Oy 121 TACGGCGATTTGATAGATACGCTGATCTTCACAGGCTTTGGCTTATATACTAC 180
Db 121 TACGGCGATTTGATAGATACGCTGATCTTCACAGGCTTTGGCTTATATACTAC 180
Oy 181 AGACATGTGAGGAAGCAGTCGACGCCAAGAGGCTCTTCAAGAGCAAAATTTGAATGA 240
Db 181 AGACATGTGAGGAAGCAGTCGACGCCAAGAGGCTCTTCAAGAGCAAAATTTGAATGA 240
Oy 241 AGTCAATTAAGATCGAATACGCAACGCGCAAAACCTTGTAAAGTCTATGGGTGGT 300
Db 241 AGTCAATTAAGATCGAATACGCAACGCGCAAAACCTTGTAAAGTCTATGGGTGGT 300
Oy 301 GGAATCGGCGCTTAATGTCTCCAGATGACCTGGAGGAAGGTCACCAAGTTGGGAAA 360
Db 301 GGAATCGGCGCTTAATGTCTCCAGATGACCTGGAGGAAGGTCACCAAGTTGGGAAA 360
Oy 361 ATCGAGATTTTAGGTTTCTCAGAAAGCAAGACAGCTTCATTGATTTATGAGATG 420
Db 361 ATCGAGATTTTAGGTTTCTCAGAAAGCAAGACAGCTTCATTGATTTATGAGATG 420
Oy 421 GATGATGCTTTACAGGCTAAGAGCATGATGAAAGGCTAAGGTTAGCTTTTTCGCT 480
Db 421 GATGATGCTTTACAGGCTAAGAGCATGATGAAAGGCTAAGGTTAGCTTTTTCGCT 480
Oy 481 GTTGATTTTCTCGGTCACAAAGCCCAAAAAGAACAAATGGGTGCTCTTACGATAC 540
Db 481 GTTGATTTTCTCGGTCACAAAGCCCAAAAAGAACAAATGGGTGCTCTTACGATAC 540
Oy 541 AGAATGGCAATATGATCAATTAACCCGATATCTCTCACTATATGAAGCTTTAAAGA 600
Db 541 AGAATGGCAATATGATCAATTAACCCGATATCTCTCACTATATGAAGCTTTAAAGA 600
Oy 601 GATGTCAGCAAGTAAGTCTGTGATGGGTCCCTCCCTACTGCTTAACAATGCAAT 660
Db 601 GATGTCAGCAAGTAAGTCTGTGATGGGTCCCTCCCTACTGCTTAACAATGCAAT 660
Oy 661 GATGAGCAAAATTTGCAAAATGCGAGTACTCTTTGGTGAATCGAGAGGTAAAAAGT 720
Db 661 GATGAGCAAAATTTGCAAAATGCGAGTACTCTTTGGTGAATCGAGAGGTAAAAAGT 720
Oy 721 TACCATCAAGAAATTTGCACTGTGAGATTTAGAGCGCGGAGAGAGTCCCAATGC 780
Db 721 TACCATCAAGAAATTTGCACTGTGAGATTTAGAGCGCGGAGAGAGTCCCAATGC 780
Oy 781 AAGGAAGGCTACAGGGAGGTTATTCATATCTAGAAATCAAAATTTATGTAATAAC 840
Db 781 AAGGAAGGCTACAGGGAGGTTATTCATATCTAGAAATCAAAATTTATGTAATAAC 840
Oy 841 GATGATGTGCTCTCTGAGCAAGACGATACTAGTTTACTGTGATGAAGCGTCAAG 900
Db 841 GATGATGTGCTCTCTGAGCAAGACGATACTAGTTTACTGTGATGAAGCGTCAAG 900
Oy 901 ACAGATATGTTCAATATGATCTTCACTGTGATCTTCCTCATCTCTGTAATTCCT 960
Db 901 ACAGATATGTTCAATATGATCTTCACTGTGATCTTCCTCATCTCTGTAATTCCT 960
Oy 961 GGGTCTATGAGGCGCTCAGAGGTACGAATGAGCGTTCAATATATGTTGTCAGAAATCAAT 1020
Db 961 GGGTCTATGAGGCGCTCAGAGGTACGAATGAGCGTTCAATATATGTTGTCAGAAATCAAT 1020
Oy 1021 GACGTTGTTGTAAGAGCAAACTGAGAGGCGCATCTGCAAAATGGAATGGAATACATC 1080
Db 1021 GACGTTGTTGTAAGAGCAAACTGAGAGGCGCATCTGCAAAATGGAATGGAATACATC 1080
Oy 1081 CCATCTCAACAGAGCACTGGAATCTCCATCTCTGCAACAAAGTACGAGGCGCTATG 1140
Db 1081 CCATCTCAACAGAGCACTGGAATCTCCATCTCTGCAACAAAGTACGAGGCGCTATG 1140
Oy 1141 AGGTCAACCCGATCTTGGGAAGATATGATCTGCTCAATGGTGGTGAAGAAATGTA 1200
Db 1141 AGGTCAACCCGATCTTGGGAAGATATGATCTGCTCAATGGTGGTGAAGAAATGTA 1200

Db 1141 AGGTCAACCCGATCTTGGGAAGATATGATCTGCTCAATGGTGGTGAAGAAATGTA 1200
Oy 1201 CGAACCAGAAAGATGATGATGAGAGGTTTACTCCAAATGGGTGTGATGAGAGTCA 1260
Db 1201 CGAACCAGAAAGATGATGATGAGAGGTTTACTCCAAATGGGTGTGATGAGAGTCA 1260
Oy 1261 TTTGTCGAGGTTGATGCTGCTAGACATCCGTGCGGCCCCCTGATTTGATACATA 1320
Db 1261 TTTGTCGAGGTTGATGCTGCTAGACATCCGTGCGGCCCCCTGATTTGATACATA 1320
Oy 1321 TGGAGAGAAATGATTTGCCAAGGGTGAAGTCCGCTGTGTTGCTGCTGTTACCTATG 1380
Db 1321 TGGAGAGAAATGATTTGCCAAGGGTGAAGTCCGCTGTGTTGCTGCTGTTACCTATG 1380
Oy 1381 GGAAGGGGATTGAATTAATCTGCTGAGAGTGTCAATTTGTCAGACAACTGATTTG 1440
Db 1381 GGAAGGGGATTGAATTAATCTGCTGAGAGTGTCAATTTGTCAGACAACTGATTTG 1440
Oy 1441 AATATGCTCGCTAAACATTAACGCGGTTGCCATTTGATGATGATGATTTTTCGTAACA 1500
Db 1441 AATATGCTCGCTAAACATTAACGCGGTTGCCATTTGATGATGATGATTTTTCGTAACA 1500
Oy 1501 GACAGGAGACAGATTTTGGCTTTACACTGAATTTCTCGGTACCTTAGCTCAAAAGAT 1560
Db 1501 GACAGGAGACAGATTTTGGCTTTACACTGAATTTCTCGGTACCTTAGCTCAAAAGAT 1560
Oy 1561 CGGCGGGGTTTGCCAAATTAAGATGATGATGATGATGATGATGATGATGATGATGAT 1620
Db 1561 CGGCGGGGTTTGCCAAATTAAGATGATGATGATGATGATGATGATGATGATGATGAT 1620
Oy 1621 TTTCTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680
Db 1621 TTTCTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680
Oy 1681 TTTACCCCGGCGACCGCTCTGTTACAGATCAATACAGCAAGAAATCTCAATCAATCT 1740
Db 1681 TTTACCCCGGCGACCGCTCTGTTACAGATCAATACAGCAAGAAATCTCAATCAATCT 1740
Oy 1741 CTGCATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1800
Db 1741 CTGCATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1800
Oy 1801 CCTAGGAAATTTACATTAAGGGGTGACAGCAAAATTTGACAGCTGCTCAAAACATCT 1860
Db 1801 CCTAGGAAATTTACATTAAGGGGTGACAGCAAAATTTGACAGCTGCTCAAAACATCT 1860
Oy 1861 GTTAGGAGGCTCTGCAATACCTATATGATGAGGCGCTCAAGCTGGGGTATGTTAACT 1920
Db 1861 GTTAGGAGGCTCTGCAATACCTATATGATGAGGCGCTCAAGCTGGGGTATGTTAACT 1920
Oy 1921 CCGAGGCTTTTAGGCACTCTGCAATATCTTCCGCAACTTCAACCTCTGCGGCT 1980
Db 1921 CCGAGGCTTTTAGGCACTCTGCAATATCTTCCGCAACTTCAACCTCTGCGGCT 1980
Oy 1981 GAGAGTCACCAACTATGTCAGAGCTTTAACAAGTTGTTTCCACAGCACATCATGTCAT 2040
Db 1981 GAGAGTCACCAACTATGTCAGAGCTTTAACAAGTTGTTTCCACAGCACATCATGTCAT 2040
Oy 2041 GGAATGCAATGAGAAACAGGCTCAAGCTTGGAAAAGGTCACAAACAGTTAT 2100
Db 2041 GGAATGCAATGAGAAACAGGCTCAAGCTTGGAAAAGGTCACAAACAGTTAT 2100
Oy 2101 GATGCGTCAATCACTCATTTCCAAACATAGGAAATCACTACAGCTGGGCACTA 2160
Db 2101 GATGCGTCAATCACTCATTTCCAAACATAGGAAATCACTACAGCTGGGCACTA 2160
Oy 2161 CTTCT 2220
Db 2161 CTTCT 2220
Oy 2221 GTTCATGCGACATCAATACAGAGCCCAATCTGTTAACTGCTCACTGCTCTCTCTCT 2280
Db 2221 GTTCATGCGACATCAATACAGAGCCCAATCTGTTAACTGCTCACTGCTCTCTCTCTCT 2280

QY	1659	GGTATATGGTGTGTTCTCAAGTTACCCCGCAGACCGTCTCTGTATACATCATACAG	1718
Db	3546	GCTATATGGTGTGTTCTCAAGTTACCCCGCAGACCGTCTCTGTATACATCATACAG	3605
QY	1719	ACAAAGATCTCAGTCCAAATCCCTGTCATTTATATGGATCAAGCCCGGGATTCACTGGCAA	1778
Db	3606	ACAAAGATCTCAGTCCAAATCCCTGTCATTTATATGGATCAAGCCCGGGATTCACTGGCAA	3665
QY	1779	TGCTAGTCAACAGTTTATATCTCTCTAGGGAAATTTACTTAAAGGGGTGCACAGAACTTT	1838
Db	3666	TGCTAGTCAACAGTTTATATCTCTCTAGGGAAATTTACTTAAAGGGGTGCACAGAACTTT	3725
QY	1839	GACAGCTGCTTCAAAACCATCTGTTAGGAGACGCTCTCGAATTAAGTACCTAATATGAAGGGCC	1898
Db	3726	GACAGCTGCTTCAAAACCATCTGTTAGGAGACGCTCTCGAATTAAGTACCTAATATGAAGGGCC	3785
QY	1899	TCAGCTGGGGTTAGTTTAACCTCCGAGACTTTTAAAGCACTGTGCATCTATTCTCCCTGC	1958
Db	3786	TCAGCTGGGGTTAGTTTAACCTCCGAGACTTTTAAAGCACTGTGCATCTATTCTCCCTGC	3845
QY	1959	AACCTCTCAACCTGTGCTCCCTGGAGAGACCAACCATATGTCAGAGACTTTCACAGTTGT	2018
Db	3846	AACCTCTCAACCTGTGCTCCCTGGAGAGACCAACCATATGTCAGAGACTTTCACAGTTGT	3905
QY	2019	TTCCACAGCAGATCAGTCCAAATGGACGTGTATACATATGAGAAAGCAGCCTCTCAAGTTGGAA	2078
Db	3906	TTCCACAGCAGATCAGTCCAAATGGACGTGTATACATATGAGAAAGCAGCCTCTCAAGTTGGAA	3965
QY	2079	AAGAGGTCCAAACAGTTCAATGATGGGTCAAAATCATCATCTTCCAAATATACGAAATCA	2138
Db	3966	AAGAGGTCCAAACAGTTCAATGATGGGTCAAAATCATCATCTTCCAAATATACGAAATCA	4025
QY	2139	GTACACTCCAGCTGGGCAACTACCTCCCTCTCCGTTACCCCTCAGCTTCAACAA	2198
Db	4026	GTACACTCCAGCTGGGCAACTACCTCCCTCTCCGTTACCCCTCAGCTTCAACAA	4085
QY	2199	CCCCAATCACTATGTGGATGATGTCATATGGCACATGCAATACCAAGCCATCTGTTAA	2258
Db	4086	CCCCAATCACTATGTGGATGATGTCATATGGCACATGCAATACCAAGCCATCTGTTAA	4145
QY	2259	CATGCGTCAGTGTCTCCGTACCAAAATATGCGTCATATTAATTTCCATGTACACTCA	2318
Db	4146	CATGCGTCAGTGTCTCCGTACCAAAATATGCGTCATATTAATTTCCATGTACACTCA	4205
QY	2319	GGGTTCTGTCAAATATCTCTGTTTCTCAAGCCATGTCAGCAATACCAACAGAGCGTC	2378
Db	4206	GGGTTCTGTCAAATATCTCTGTTTCTCAAGCCATGTCAGCAATACCAACAGAGCGTC	4265
QY	2379	CATGCCAAACCAAACTATGGTCCAAATTCCAAGTTATCAGCAAGCTAATTTTATGAGCGT	2438
Db	4266	CATGCCAAACCAAACTATGGTCCAAATTCCAAGTTATCAGCAAGCTAATTTTATGAGCGT	4325
QY	2439	AACACAAATTCAGGACAGACAGATTAAACCTTCCCAATTTCAAGCTGCCATGCACACCC	2498
Db	4326	AACACAAATTCAGGACAGACAGATTAAACCTTCCCAATTTCAAGCTGCCATGCACACCC	4385
QY	2499	AGCAGATTAAGGCAAAATTTAAGCGCAAAACCAAGACATACAGATTGACAGCTATGATCTC	2558
Db	4386	AGCAGATTAAGGCAAAATTTAAGCGCAAAACCAAGACATACAGATTGACAGCTATGATCTC	4445
QY	2559	TGGGGATGTCAGGGTCAACAGATGGGGAGGTGATTAAGATACAGATACAGATCAAC	2618
Db	4446	TGGGGATGTCAGGGTCAACAGATGGGGAGGTGATTAAGATACAGATACAGATCAAC	4505
QY	2619	ACTACATTTTGGACGAAACCTTTCTTCCAGATATACAGCAAGAAACAGACAGACATGCTTC	2678
Db	4506	ACTACATTTTGGACGCAAACTTTCTTCTCCAGATATACAGCAAGAAACAGACAGATGCTTC	4565
QY	2679	AGGTACTCTCGGCTGGAGAGGGGCTTTGA	2706
Db	4566	AGGTACTCTCGGCTGGAGAGGGGCTTTGA	4593

	RESULT 3	
	US-09-920-705-4/C	
	; Sequence 4, Application US/09920705	
	; Publication No. US20030079252A1	
	; GENERAL INFORMATION:	
	; APPLICANT: Amasilo, Richard M.	
	; APPLICANT: Schomburg, Fritz M.	
	; APPLICANT: Michaels, Scott D.	
	; APPLICANT: Patton, David	
	; TITLE OF INVENTION: Floral Induction Gene	
	; FILE REFERENCE: 960296.97214	
	; CURRENT APPLICATION NUMBER: US/09/920,705	
	; CURRENT FILING DATE: 2001-08-02	
	; NUMBER OF SEQ ID NOS: 6	
	; SOFTWARE: PatentIn Ver. 2.1	
	; SEQ ID NO 4	
	; LENGTH: 801	
	; TYPE: DNA	
	; ORGANISM: Arabidopsis thaliana	
	; FEATURE:	
	; OTHER INFORMATION: FPA antisense fragment	
	US-09-920-705-4	
	Query Match	29.5%; Score 797.8; DB 9; Length 801;
	Best Local Similarity 99.8%; Pred. No.2,1e-245;	
	Matches 799; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	
OY	587 AAGACTTAAAGAGATGTCACGCCAAGTAAGTTCTGTGATTGGTTCCTCCTACTG	646
DB	801 .AAGACTTAAAGAGATGTCACGCCAAGTAAGTTCTGTGATTGGTTCCTCCTACTG	742
OY	647 CTACCAATGCAATGATAGCAAAATTCGCACAATGCATATATCTTTGGTAGATCG	706
DB	741 CTACCAATGCAATGATAGCAAAATTCGCACAATGCATATATCTTTGGTAGATCG	682
OY	707 AGAGGGTAAAAAGTTACCACATCAAGGAATTTTGCACTTGTGAGTTTAGAGCGCGAGG	766
DB	681 AGAGGGTAAAAAGTTACCACATCAAGGAATTTTGCACTTGTGAGTTTAGAGCGCGAGG	622
OY	767 AAGCTCCGCATGCAAGAAGGCTCACAGGGGAGTTATTCAATATCTTGAATCAA	826
DB	621 AAGCTCCGCATGCAAGAAGGCTCACAGGGGAGTTATTCAATATCTTGAATCAA	562
OY	827 TTATGTACTCAAGCATGATTGCTCCTGAGCAAGACGATACTGTTTACTCTG	886
DB	561 TTATGTACTCAAGCATGATTGCTCCTGAGCAAGACGATACTGTTTACTCTG	502
OY	887 TGAACGGTCAAGCAGATATGTTCAATTAATGATCCTTCAATGTGTATCTTCTCATT	946
DB	501 TGAACGGTCAAGCAGATATGTTCAATTAATGATCCTTCAATGTGTATCTTCTCATT	442
OY	947 CTACTGCAATTCCTGGGTCATGTGAGGCCCTCAGAGGATACCAATGAGGGTTCATATATG	1006
DB	441 CTACTGCAATTCCTGGGTCATGTGAGGCCCTCAGAGGATACCAATGAGGGTTCATATATG	382
OY	1007 GTGCAGATACATGACGTTGTGTGAAGAGCCAAAAGTGAAGAGGCCATCTGCCAATG	1066
DB	381 GTGCAGATACATGACGTTGTGTGAAGAGCCAAAAGTGAAGAGGCCATCTGCCAATG	322
OY	1067 GAATGTGAATTCCTCCCATCTCCACAGAGACCTGGAATTCCTCCCATCTCCGACAAAGTA	1126
DB	321 GAATGTGAATTCCTCCCATCTCCACAGAGACCTGGAATTCCTCCCATCTCCGACAAAGTA	262
OY	1127 CGAGGCGCCCTATGAGGTCAAAACCCCGATTCCTTGGGAAGATATGATCCTGCTCAGTTGG	1186
DB	261 CGAGGCGCCCTATGAGGTCAAAACCCCGATTCCTTGGGAAGATATGATCCTGCTCAGTTGG	202
OY	1187 TCAGAGAAATGAAACGAAACGAAAGAGATGATGACGAGGTTTTACTCCATGGGGT	1246
DB	201 TCAGAGAAATGAAACGAAACGAAAGAGATGATGACGAGGTTTTACTCCATGGGGT	142
OY	1247 TCATGTAGAGATCATTTGGTCGAGGGTTCAAGTTGTGTCTAACCATCCGAGGGCCCCCTG	1306

Accession	Sequence	Position
Dd	141 TCATGAGAGAGTCATTGGTCGAGGTTCACTGCTGTAACACATACCGTGCCCCCTG	82
OY	1307 ATTTCGATCACAATGAGAGGAATGATTCGCCAAGGGTGAACTCCCGTCTGTTGTGCTC	1366
Dd	81 ATTTCGATCACAATGAGAGGAATGATTCGCCAAGGGTGGAACCTCCGCTGTGTGTCTC	22
OY	1367 GTTGTGACCTATGGGAAAGG 1387	
Dd	21 GTTGTGACCTATGGGAAAGG 1	

```

RESULT 4
US-09-920-705-5
Sequence 5, Application US/09920705
Publication No. US20030079252A1
GENERAL INFORMATION:
APPLICANT: Amasino, Richard M.
APPLICANT: Schomberg, Fritz M.
APPLICANT: Michaels, Scott D.
APPLICANT: Patton, David
TITLE OF INVENTION: Floral Induction Gene
FILE REFERENCE: 960296.97214
CURRENT APPLICATION NUMBER: US/09/920,705
CURRENT FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 6
SOFTWARE: Patentl Ver. 2.1
SEQ ID NO 5
LENGTH: 349
TYPE: DNA
ORGANISM: Arabidopsis thaliana
FEATURE:
OTHER INFORMATION: Portion of FPA coding region
US-09-920-705-5

```

Query Match	10.0%	Score 271.4	DB 9	Length 349
Best Local Similarity	94.6%	Pred. No 1.1e-77		
Matches	281	Conservative	0	Mismatches 16; Indels 0; Gaps 0
QY	1	ATGCGCTATCTATGAAAGCCATTCGAGCCGATGATTCGGTTCACGTCAAAACAACTTT	60	
Db	1	ATGCGCTATCTATGAAAGCCATTCGAGCCGATGATTCGGTTCACGTCAAAACAACTTT	60	
QY	61	TGGTTCGGTACCTTAACGCGCGAGACGACAGTATGATCTGACCAGATGTTGTTGGAGA	120	
Db	61	TGGTTCGGTACCTTAACGCGCGAGACGACAGTATGATCTGACCAGATGTTGTTGGAGA	120	
QY	121	TACGGCGATTTGTTATAGATCAACGGTGATTTCTTCACGAGGCTTTGCGTTTATATATATAC	180	
Db	121	TACGGCGATTTGTTATAGATCAACGGTGATTTCTTCACGAGGCTTTGCGTTTATATATATAC	180	
QY	181	AGACATGCGAGGAGAACGTCGACAGCCAAAGGGCTCTTCAAGGAGCAAAATTTGAAATGGA	240	
Db	181	AGACATGCGAGGAGAACGTCGACAGCCAAAGGGCTCTTCAAGGAGCAAAATTTGAAATGGA	240	
QY	241	AGTCAAAATTAAGATCGAATACGACAGACCGGCCAAAACCTTGTAAGAGTCATATGGGTG	297	
Db	241	AGTCAAAATTAAGATCGAATACGACAGACCGGGTTGTTCTTATATATCTATCTCGTTGG	297	

```

RESULT 5
US-09-920-705-6
: Sequence 6, Application US/09920705
: Publication No. US20030079252A1
: GENERAL INFORMATION:
: APPLICANT: Amasino, Richard M.
: APPLICANT: Schomburg, Fritz M.
: APPLICANT: Michaels, Scott D.
: APPLICANT: Patton, David
: TITLE OF INVENTION: Floral Induction Gene
: FILE REFERENCE: 960296.97314
: CURRENT APPLICATION NUMBER: US/09/920,705
: CURRENT FILING DATE: 2001-08-02

```

```

: NUMBER OF SEQ ID NOS: 6
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 6
: LENGTH: 3715
: TYPE: DNA
: ORGANISM: Arabidopsis thaliana
: FEATURE:
: OTHER INFORMATION: FPA promoter plus intron
US-09-920-705-6

```

	Query Match	10.0%	Score 271.4	DB: 9	Length 3715
	Best Local Similarity	94.6%	Pred. No. 6.8e-77		
	Matches 281	Conservative 0	Mismatches 16	Indels 0	Gaps 0
Oy	1	ATGGCGTTATCTATGAAGCCATTGCAGAGCCGATGATTTCCGGTTTCAGTCACCAATCTT	60		
Db	1832	ATGGCGTTATCTATGAAGCCATTGCAGAGCCGATGATTTCCGGTTTCAGTCACCAATCTT	1891		
Oy	61	TGGGTGGGTACCCATCAAGCCGGAGACGACAGTCAGATCAGTCGATCCGAGTGTGGAGA	120		
Db	1892	TGGGTGGGTACCCATCAAGCCGGAGACGACAGTCAGATCAGATCTGACCCGATTTTGGAGA	1951		
Oy	121	TACGGCGATATTGATAGAAATCAGCGTGTATTCTTACAGAGGCTTTGCGTTTATATCTAC	180		
Db	1952	TACGGCGATATTGATAGAAATCAGCGTGTATTCTTACAGAGGCTTTGCGTTTATATCTAC	2011		
Oy	181	AGCATGTGGGAGGAGGAGCGTGCAGAGCCAAAGAGCTCTTCAAGAGGCAAAATTTAAATGA	240		
Db	2012	AGCATGTGGGAGGAGGAGCGTGCAGAGCCAAAGAGCTCTTCAAGAGGCAAAATTTAAATGA	2071		
Oy	241	AGTCAAAATTAAGATCGAATTCGACAGACCGGCAAAACCTTGTAAAGTCATGGGCTG	297		
Db	2072	AGTCAAAATTAAGATCGAATTCGACAGACCGGTTTGTCTTATTCATATCTTCGCTTG	2128		

```

RESULT 6
US-09-770-444-972
? Sequence 972: Application US/09770444
? Patent No. US20020023280A1
? GENERAL INFORMATION:
?   APPLICANT: Goriach, Jörn
?   APPLICANT: An, Yong-Qiang
?   APPLICANT: Hamilton, Carol M.
?   APPLICANT: Price, Jennifer L.
?   APPLICANT: Raines, Tracy M.
?   APPLICANT: Yu, Yang
?   APPLICANT: Rameeka, Joshua G.
?   APPLICANT: Page, Amy
?   APPLICANT: Matthew, Abraham V.
?   APPLICANT: Ledford, Brooke L.
?   APPLICANT: Woessner, Jeffrey P.
?   APPLICANT: Haas, William David
?   APPLICANT: Garcia, Carlos A.
?   APPLICANT: Kricker, Maja
?   APPLICANT: Slader, Ted
?   APPLICANT: Davis, Keith R.
?   APPLICANT: Allen, Keith
?   APPLICANT: Hoffman, Neil
?   APPLICANT: Hubban, Patrick
?   TITLE OF INVENTION: Expressed sequences of Arabidopsis
?   TITLE OF INVENTION: thaliana
?   FILE REFERENCE: 2027 (PARA-016PRV)
?   CURRENT APPLICATION NUMBER: US/09/770,444
?   CURRENT FILING DATE: 2001-01-26
?   PRIOR APPLICATION NUMBER: 60/178,502
?   PRIOR FILING DATE: 2000-01-27
?   NUMBER OF SEQ ID NOS: 999
?   SOFTWARE: FastSeq for Windows Version 4.0
?   SEQ ID NO 972
?   LENGTH: 438
?   TYPE: DNA
?   ORGANISM: Arabidopsis thaliana
US-09-770-444-972

```

Query Match 4.2%; Score 113.2; DB 10; Length 438;
Best Local Similarity 61.6%; Pred. No. 6.7e-26;
Matches 181; Conservative 0; Mismatches 113; Indels 0; Gaps 0;

QY 1390 ATTGAACCTAAACGCTGAGAGGTGCTCAATTGTCACGACAACTGATTTGAATATGCTC 1449
DB 1 ATGACATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 60

QY 1450 GCTAAACATACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1509
DB 61 GCTAAACATACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120

QY 1510 GAAGATTTTGGCTTTACCTGAAATTTCTCCGCTACCTTACCTCAAAAGATGGGCGGCT 1569
DB 121 GCTGATATCGCTTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 180

QY 1570 GTTGCAAAATAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1629
DB 181 GTTCTTAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 240

QY 1630 GATGATCTCAAGTACCGCTGCAAGAACGCTATATGCTGCTGCTGCTGCTGCTGCTGCTG 1683
DB 241 AAGTACTTAAGTTCCTCGGAACCTAAGCATCTCTGAGTATTTCTACGTTTA 294

RESULT 7
US-09-815-343-1381
; Sequence 1381, Application US/09815343
; Patent No. US2001005596A1
; GENERAL INFORMATION:
; APPLICANT: Meagher, Madeleine
; APPLICANT: Xu, Jiangchun
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; FILE REFERENCE: 210121.504
; CURRENT APPLICATION NUMBER: US/09/815.343
; CURRENT FILING DATE: 2001-03-22
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1381
; LENGTH: 256
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(256)
; OTHER INFORMATION: n = A,T,C or G
US-09-815-343-1381

Query Match 1.6%; Score 42.8; DB 10; Length 256;
Best Local Similarity 59.7%; Pred. No. 0.0045;
Matches 71; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 151 TCTTACAGAGCTTTGGCTTTATATCTACACATGTGAGAGAGTGGAGCCAAA 210
DB 89 TCTTCCAAAGGCTTTGGCTTTTGTAGACTTAAACAGTACGAGGATGCCAAAGTCCCAAN 148

QY 211 GAGGCTTCTCAAGAGCAAAATTTGAATGAAGTCAAATTAAGTATGCAATAGCAGCAGCC 269
DB 149 GAGGCCATGGAAGACGCTGAATTTGATGAAATTAAGTTACTTGGAGCTGGGCCAAACC 207

RESULT 8
US-09-917-800A-1567
; Sequence 1567, Application US/09917800A
; Patent No. US20020119462A1
; GENERAL INFORMATION:
; APPLICANT: Mendrick, Donna
; APPLICANT: Porter, Mark
; APPLICANT: Johnson, Kory
; APPLICANT: Castle, Arthur

APPLICANT: Elashoff, Michael
APPLICANT: Gene Logic, Inc.
TITLE OF INVENTION: Molecular Toxicology Modeling
FILE REFERENCE: 44921-5038-US
CURRENT APPLICATION NUMBER: US/09/917.800A
PRIOR FILING DATE: 2001-07-31
PRIOR APPLICATION NUMBER: US 60/222,040
PRIOR FILING DATE: 2000-07-31
PRIOR APPLICATION NUMBER: US 60/222,880
PRIOR FILING DATE: 2000-11-02
PRIOR APPLICATION NUMBER: US 60/290,029
PRIOR FILING DATE: 2001-05-11
PRIOR APPLICATION NUMBER: US 60/290,645
PRIOR FILING DATE: 2001-05-15
PRIOR APPLICATION NUMBER: US 60/292,336
PRIOR FILING DATE: 2001-05-22
PRIOR APPLICATION NUMBER: US 60/295,798
PRIOR FILING DATE: 2001-06-06
PRIOR APPLICATION NUMBER: US 60/297,457
PRIOR FILING DATE: 2001-06-13
PRIOR APPLICATION NUMBER: US 60/298,884
PRIOR FILING DATE: 2001-06-19
PRIOR APPLICATION NUMBER: US 60/303,459
PRIOR FILING DATE: 2001-07-09
NUMBER OF SEQ ID NOS: 1740
SOFTWARE: Patentln Ver. 2.1
SEQ ID NO 1567
LENGTH: 2142
TYPE: DNA
ORGANISM: Rattus norvegicus
FEATURE:
OTHER INFORMATION: Genbank Accession No. US20020119462A1 NM_012749
US-09-917-800A-1567

Query Match 1.5%; Score 41.6; DB 10; Length 2142;
Best Local Similarity 59.2%; Pred. No. 0.0053;
Matches 71; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 150 TCTTACAGAGCTTTGGCTTTTATATCTACACATGTGAGAGAGTGGAGCAGTCGACGCA 209
DB 1830 TCTTCTAAAGGCTTTGGCTTTTGTAGACTTAAATAGAGAGAGATGCCAAAGCTGCCAA 1889

QY 210 AGAGCTCTTCAAGAGCAAAATTTGAATGAGACTCAATTAAGATGATGATGATGATGATGATG 269
DB 1890 GAGGCCATGGAAGATGAGCAAAATTTGATGAAACCAAAAGTTACCTGGAGCTGGGCCAAACC 1949

RESULT 9
US-10-123-155-438
; Sequence 438, Application US/10123155
; Publication No. US20030068794A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Oiang
; APPLICANT: Geriltsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C30
; CURRENT APPLICATION NUMBER: US/10/123.155

US-RESULT 10
US-10-062-254-287
Sequence 287, Application US/10062254
Patent No. US20020138862A1
GENERAL INFORMATION:
APPLICANT: Cahoon, Edgar B
APPLICANT: Cahoon, Rebecca E
APPLICANT: Falco, Saverio Carl
APPLICANT: Fang, Yiyen
APPLICANT: Hartke, Sabine S.
APPLICANT: Lee, Jian-Ming
APPLICANT: Li, Zhongsen
APPLICANT: Miao, Guo-Hua
APPLICANT: Morgante, Michele
APPLICANT: Niu, Xiping
APPLICANT: Odell, Joan
APPLICANT: Rafalski, Antoni
APPLICANT: Sakai, Hajime
APPLICANT: Zheng, Peizhong

QY	290	TATGGGTGGTGAATCGGCGCTTATGTCCTCCAAAGATGACCTGAGAGAGATTCAGCA	349
Db	503	TATTTTGTGTGGCTTGACCCCAATGTCACTGGAAGATGCTTGAAACAAAGTGTTCCTC	562
QY	350	AGTTTGGAAAAATCGAGATTTTAACTTCTTCAGAAAGCAGACAGCTTTCATGATT	409
Db	563	CAATATGGGAAGTTGTCATGTCAAGATGCCGTGTGGGAAGAGATGTGGCTTGTACGT	622
QY	410	ATTATGAGATGATGATGCTTTACAGGCTTAAGACCATGAATGAAAGCCTATGGTGCTA	469
Db	623	ATGCTAAACGCGCTTCTGCTGAGCAAGCTCTCAATTCCTCAAGAGGACCTTGGTGGTG	682
QY	470	G 470	
Db	683	G 683	

US-RJOUT-11-155-358
Sequence 358, Application US/10123155
Publication No. US20030068794/1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Bereshki, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Andrew
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tunas, Daniel

RESULT 12
US-09-878-574-4075
; Sequence 4075, Application US/09878574
; Patent No. US20020110548A1

RESULT 13
 US-10-062-254-275
 Patent No. US20020138882A1
 GENERAL INFORMATION:
 APPLICANT: Cahoon, Edgar B
 APPLICANT: Cahoon, Rebecca E
 APPLICANT: Falco, Saverio Carl
 APPLICANT: Fang, Yiwen
 APPLICANT: Hantke, Sabine S.
 APPLICANT: Lee, Jian-Ming
 APPLICANT: Li, Zhongsen
 APPLICANT: Miao, Guo-Hua
 APPLICANT: Morgante, Michele
 APPLICANT: Niu, Xiping
 APPLICANT: Odell, Joan
 APPLICANT: Rafalski, Antoni
 APPLICANT: Sakai, Hajime
 APPLICANT: Zheng, Peizhong
 APPLICANT: Zhu, Qun
 TITLE OF INVENTION: Polynucleotides Encoding Proteins Involved in Plant Metabolism
 FILE REFERENCE:
 CURRENT APPLICATION NUMBER: US/10/062,254
 CURRENT FILING DATE: 2002-02-01
 PRIOR APPLICATION NUMBER: 09/630,346
 PRIOR FILING DATE: 2000-07-28
 PRIOR APPLICATION NUMBER: 60/146511
 PRIOR FILING DATE: 1999-07-30
 PRIOR APPLICATION NUMBER: 60/156006
 PRIOR FILING DATE: 1999-09-23
 PRIOR APPLICATION NUMBER: 60/156899
 PRIOR FILING DATE: 1999-09-30
 PRIOR APPLICATION NUMBER: 60/157287
 PRIOR FILING DATE: 1999-10-01
 PRIOR APPLICATION NUMBER: 60/169767
 PRIOR FILING DATE: 1999-12-09
 PRIOR APPLICATION NUMBER: 60/171054
 PRIOR FILING DATE: 1999-12-16
 PRIOR APPLICATION NUMBER: 60/172958
 PRIOR FILING DATE: 1999-12-21
 PRIOR APPLICATION NUMBER: 60/171515

;; PRIOR FILING DATE: 1999-12-22
;; PRIOR APPLICATION NUMBER: 60/173535
;; PRIOR FILING DATE: 1999-12-29
;; NUMBER OF SEQ ID NOS: 375
;; SOFTWARE: Microsoft Office 97
;; SEQ ID NO 275
;; LENGTH: 1537
;; TYPE: DNA
;; ORGANISM: Glycine max
US-10-062-254-275

Query Match 1.4%; Score 37.2; DB 12; Length 1537;
Best Local Similarity 61.2%; Pred. No. 1.2;
Matches 60; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

OY 282 TTACAGTCTGATGGTGGTGAATCGCCCTAATGTCTCCAGAGTGAAGTGAAGA 341
DB 879 TACACTATTTTGTGTGTAATTTGGATCTAATGTCTACATGATCATCTGAGCAAGT 938
OY 342 GTTCAGCAAGTTTGGGAAATCGAGATTATTAGCTTC 379
DB 939 TTTCAGCCAGTATGAGATTAGTGCATGTGAAGATTC 976

RESULT 14
US-10-197-666A-77
;; Sequence 77, Application US/10197666A
;; Publication No. US20030092037A1
;; GENERAL INFORMATION:
;; APPLICANT: ASAH KASEI KABUSIKI KAISYA
;; TITLE OF INVENTION: ELKI phosphorylation related gene
;; FILE REFERENCE: PH-1548US
;; CURRENT APPLICATION NUMBER: US/10/197,666A
;; PRIOR FILING DATE: 2002-11-18
;; PRIOR APPLICATION NUMBER: JP 2001-218204
;; PRIOR FILING DATE: 2001-07-18
;; PRIOR APPLICATION NUMBER: JP 2001-263450
;; PRIOR FILING DATE: 2001-08-31
;; PRIOR APPLICATION NUMBER: JP 2002-012176
;; PRIOR FILING DATE: 2002-01-21
;; PRIOR APPLICATION NUMBER: US 60/305,884
;; PRIOR FILING DATE: 2001-07-18
;; PRIOR APPLICATION NUMBER: US 60/316,304
;; PRIOR FILING DATE: 2001-09-04
;; PRIOR APPLICATION NUMBER: US 60/350,027
;; PRIOR FILING DATE: 2002-01-23
;; NUMBER OF SEQ ID NOS: 156
;; SOFTWARE: Patentln Ver. 2.1
;; SEQ ID NO 77
;; LENGTH: 1846
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: (128)..(877)
US-10-197-666A-77

Query Match 1.4%; Score 37; DB 9; Length 1846;
Best Local Similarity 54.9%; Pred. No. 1.6;
Matches 73; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

OY 153 TTCACGAGGCTTTGGTTTATATCTACAGACATGTGAGAGCGACGCCAAGA 212
DB 484 TTCAAGAGGATTTGCTTTGTAATTTTGAATAATGTAGATGCGCAAGAGCTAAGA 543
OY 213 GGCCTTTCAGAGCAAAATTTGAATGAATGCAATTAAGATCGAATACGACGACCGGC 272
DB 544 ACGTGCCAATGGAATGAGCTTGATGAGCGGTAGATCAGAGTTTCTCTATACAAA 603
OY 273 AAAACCTTGTAG 285
DB 604 AAGACCACTACG 616

RESULT 15
US-10-197-666A-79
;; Sequence 79, Application US/10197666A
;; Publication No. US20030092037A1
;; GENERAL INFORMATION:
;; APPLICANT: ASAH KASEI KABUSIKI KAISYA
;; TITLE OF INVENTION: ELKI phosphorylation related gene
;; FILE REFERENCE: PH-1548US
;; CURRENT APPLICATION NUMBER: US/10/197,666A
;; PRIOR FILING DATE: 2002-11-18
;; PRIOR APPLICATION NUMBER: JP 2001-218204
;; PRIOR FILING DATE: 2001-07-18
;; PRIOR APPLICATION NUMBER: JP 2001-263450
;; PRIOR FILING DATE: 2001-08-31
;; PRIOR APPLICATION NUMBER: JP 2002-012176
;; PRIOR FILING DATE: 2002-01-21
;; PRIOR APPLICATION NUMBER: US 60/305,884
;; PRIOR FILING DATE: 2001-07-18
;; PRIOR APPLICATION NUMBER: US 60/316,304
;; PRIOR FILING DATE: 2001-09-04
;; PRIOR APPLICATION NUMBER: US 60/350,027
;; PRIOR FILING DATE: 2002-01-23
;; NUMBER OF SEQ ID NOS: 156
;; SOFTWARE: Patentln Ver. 2.1
;; SEQ ID NO 79
;; LENGTH: 1964
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: (130)..(993)
;; FEATURE:
;; NAME/KEY: unsure
;; LOCATION: 38
;; OTHER INFORMATION: n=A or C or G or T
US-10-197-666A-79

Query Match 1.4%; Score 37; DB 9; Length 1964;
Best Local Similarity 54.9%; Pred. No. 1.6;
Matches 73; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

OY 153 TTCACGAGGCTTTGGTTTATATCTACAGACATGTGAGAGCGACGCCAAGA 212
DB 600 TTCAAGAGGATTTGCTTTGTAATTTTGAATAATGTAGATGCGCAAGAGCTAAGA 659
OY 213 GGCCTTTCAGAGCAAAATTTGAATGAATGCAATTAAGATCGAATACGACGACCGGC 272
DB 660 ACGTGCCAATGGAATGAGCTTGATGAGCGGTAGATCAGAGTTTCTCTATACAAA 719
OY 273 AAAACCTTGTAG 285
DB 720 AAGACCACTACG 732

Search completed: June 19, 2003, 07:32:24
Job time : 933 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 19, 2003, 01:11:43 ; Search time 128 Seconds
(without alignments)
6483.338 Million cell updates/sec

Title: US-09-920-705-2

Perfect score: 2706
Sequence: 1 atggcgtatctatgaagcc.....cgcgtgacagggccttga 2706Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCRTUS.COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfilest.seq.*

* Pred. NO. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Query Length	DB ID	Description
C 1	47.6	1.8	7218	1 US-08-232-463-14	Sequence 14, Appl
C 2	35.4	1.3	3489	2 US-08-728-323A-1	Sequence 1, Appl
C 3	35.4	1.3	3489	4 US-09-298-568-1	Sequence 1, Appl
C 4	35.4	1.3	32207	2 US-08-770-379-20	Sequence 20, Appl
C 5	35.4	1.3	32207	4 US-08-757-669A-20	Sequence 20, Appl
C 6	35.4	1.3	32207	4 US-09-230-371A-20	Sequence 20, Appl
C 7	34.8	1.3	2406	4 US-09-632-098-5	Sequence 5, Appl
C 8	34.8	1.3	2439	4 US-09-632-098-6	Sequence 6, Appl
C 9	34.6	1.3	561	2 US-08-589-711-3	Sequence 3, Appl
C 10	34.6	1.3	561	4 US-09-117-257-22	Sequence 22, Appl
C 11	34.6	1.3	561	4 US-09-117-257-24	Sequence 24, Appl
C 12	34.6	1.3	561	4 US-09-221-938-3	Sequence 8, Appl
C 13	34.6	1.3	561	4 US-08-945-476-8	Sequence 8, Appl
C 14	34.6	1.3	561	4 US-08-945-476-22	Sequence 22, Appl
C 15	34.6	1.3	561	4 US-08-945-476-24	Sequence 22, Appl
C 16	34.6	1.3	561	4 US-08-945-476-24	Sequence 22, Appl
C 17	34.6	1.3	561	4 US-08-945-476-24	Sequence 22, Appl
C 18	34.6	1.3	561	4 US-08-945-476-24	Sequence 22, Appl
C 19	34.6	1.3	561	4 US-08-945-476-24	Sequence 22, Appl
C 20	34.6	1.3	561	4 US-08-945-476-24	Sequence 22, Appl
C 21	34.6	1.3	561	4 US-08-945-476-24	Sequence 22, Appl
C 22	34.6	1.3	561	4 US-08-945-476-24	Sequence 22, Appl
C 23	34.6	1.3	561	4 US-08-945-476-24	Sequence 22, Appl
C 24	34.6	1.3	561	4 US-08-945-476-24	Sequence 22, Appl
C 25	34.6	1.3	561	4 US-08-945-476-24	Sequence 22, Appl
C 26	34.6	1.3	561	4 US-08-945-476-24	Sequence 22, Appl
C 27	34.6	1.3	561	4 US-08-945-476-24	Sequence 22, Appl

C 28	34.6	1.3	2656	4 US-09-117-257-7	Sequence 7, Appl
C 29	34.6	1.3	2656	4 US-09-489-352-7	Sequence 7, Appl
C 30	34.2	1.3	2849	4 US-09-221-017B-990	Sequence 990, App
C 31	34	1.3	1851	3 US-09-042-426-3	Sequence 3, Appl
C 32	34	1.3	1851	4 US-09-291-238-3	Sequence 3, Appl
C 33	34	1.3	1851	4 US-09-330-760-3	Sequence 3, Appl
C 34	34	1.3	1851	4 US-09-328-473-3	Sequence 3, Appl
C 35	34	1.3	1851	4 US-09-330-737-3	Sequence 3, Appl
C 36	34	1.3	1851	4 US-09-329-169-3	Sequence 3, Appl
C 37	34	1.3	1851	4 US-09-330-714A-3	Sequence 3, Appl
C 38	34	1.3	1851	4 US-09-328-826-3	Sequence 3, Appl
C 39	34	1.3	7378	3 US-09-042-426-9	Sequence 9, Appl
C 40	34	1.3	7378	4 US-09-291-228-9	Sequence 9, Appl
C 41	34	1.3	7378	4 US-09-330-760-9	Sequence 9, Appl
C 42	34	1.3	7378	4 US-09-328-473-9	Sequence 9, Appl
C 43	34	1.3	7378	4 US-09-330-737-9	Sequence 9, Appl
C 44	34	1.3	7378	4 US-09-329-169-9	Sequence 9, Appl
C 45	34	1.3	7378	4 US-09-330-714A-9	Sequence 9, Appl

ALIGNMENTS

RESULT 1
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEFFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOMLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232.463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935.313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: PTZgpt-F1s
; US-08-232-463-14
Query Match 1.8%, Score 47.6, DB 1, Length 7218;

Best Local Similarity 6.8%; Pred. No. 0.00015;
Matches 29; Conservative 214; Mismatches 183; Indels 0; Gaps 0;

```

OY      132 TGAATAGATCACGGGTATTCTTCCACGAGCGCTTGCGTTTATTAATCAAGACATCNGGA   191
        || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1496 TGTAGGCATCACTGTAATTACCTATCTATGCCAAGTAGTTAAAGATGAAGAATTGGT   1437
OY      192 GGAAGCAGCTCGCACCACAAGAGCGCTCTTCAAGAGACAATTGATGAGAGCTCAATTAA   251
        : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1436 ACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR   1377
OY      252 GATGGAATAGCCAGCAGCCGGCAAACCTTGTAAGACTATANGCGTGGTCGAATGGCCC   311
        : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1376 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR   1317
OY      312 TAATGTCCTCAAGGATGACCTGCGAGGAAGATGCAGAACTTGGAATAATCGAGGATT   371
        : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1316 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR   1257
OY      372 TAGCTTCTCAGAACGCCAAGACACGCTTCATTGATTATTATGAGATGATGATCTT   431
        : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1356 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR   1197
OY      432 ACAGGCTAAGAGCATGATGGAAGCGCTATGGGAGTAGCTTTTGGGTGTTGATTTCT   491
        : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1196 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR   1137
OY      492 CCGGTCAACAGCGCCAATAAAAAAGACAATGGGCTGCGCTCTTACATAACAGAAATGGCA   551
        : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1136 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR   1077
OY      552 TATGAA 557
        : : : : :
Db      1076 RRRRRR 1071

```

RESULT 2

US-08-128-323A-1/c
; Sequence 1, Application US/08728323A

GENERAL INFORMATION:
APPLICANT: Chang, Yuan
APPLICANT: Bohenzky, Roy A.
APPLICANT: Russo, James J.
APPLICANT: Edelman, Isidore S.
APPLICANT: Moore, Patrick S.
TITLE OF INVENTION: Immediate Early Protein From Kaposi's
TITLE OF INVENTION: Sarcoma-Associated Herpesvirus, DNA
TITLE OF INVENTION: Encoding Same And Uses Thereof
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30C
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/728,323A

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:

```

:      LENGTH: 3489 base pairs
:      TYPE: nucleic acid
:      STRANDEDNESS: single
:      TOPOLOGY: linear
:      MOLECULE TYPE: DNA (genomic)
:      FEATURE:
:      NAME/KEY: CDS
:      LOCATION: 1..3489
:      US-08-728-323A-1

```

Query Match	1.3%;	Score 35.4;	DB 2;	Length 3489;
Best Local Similarity	44.1%;	Pred. No. 0.92;		
Matches 147;	Conservative 0;	Mismatches 186;	Indels 0;	Gaps 0

[illegible]

RESULT 3

US-09-298-568-1/c
; Sequence 1, Application US/09298568
Patent No. 533370

```

GENERAL INFORMATION:
APPLICANT: Kieff, Elliott D.
APPLICANT: Ballestas, Mary E.
APPLICANT: Kaye, Kenneth M.
TITLE OF INVENTION: RHADINO VIRUS LANA ACTS IN TRANS ON A UNIT OF RHADINO
TITLE OF INVENTION: VIRUS DNA TO MEDIMATE EFFICIENT EPISODE PERSISTENCE
FILE REFERENCE: 16412-10001R
CURRENT APPLICATION NUMBER: US/09/298,568
CURRENT FILING DATE: 1999-04-21
EARLIER APPLICATION NUMBER: US 60/109,422
EARLIER FILING DATE: 1998-11-19
NUMBER OF SEO ID NOS: 3
SOFTWARE: Patentln Ver. 2.0
SEO ID NO 1
LENGTH: 3489
TYPE: DNA
ORANISM: Kaposi's sarcoma-associated herpesvirus
US-09-298-568-1

```

Query Match	1.38;	Score 35.4;	DB 4;	Length 3489;
Best Local Similarity	44.18;	Pred. No. 0.92;		
Matches 147;	Conservative	0;	Mismatches 186;	Indels 0;
			Gaps	0;

[illegible]

[illegible]

```

1  RESULT 8
2  US-09-632-098-6
3  ; Sequence 6, Application US/09632098
4  ; Patent No. 6420154
5  ; GENERAL INFORMATION:
6  ; APPLICANT: Sheppard, Paul O.
7  ; APPLICANT: Bairdard, Nand
8  ; APPLICANT: Bishop, Paul D.
9  ; TITLE OF INVENTION: MAMMALIAN ADHESION PROTEASE PEPTIDES
10 ; FILE REFERENCE: 99-39
11 ; CURRENT APPLICATION NUMBER: US/09/632,098
12 ; CURRENT FILING DATE: 2000-08-02
13 ; NUMBER OF SEQ ID NOS: 26
14 ; SOFTWARE: FastSeq for Windows Version 3.0
15 ; SEQ ID NO 6
16 ; LENGTH: 2439
17 ; TYPE: DNA
18 ; ORGANISM: Artificial Sequence
19 ; FEATURE:
20 ; OTHER INFORMATION: Degenerate sequence
21 ; NAME/KEY: misc_feature
22 ; LOCATION: (1)...(2439)
23 ; OTHER INFORMATION: n = A,T,C or G
24 ; US-09-632-098-6

```


FILE REFERENCE: 4210.000500
CURRENT APPLICATION NUMBER: US/09/117,257
CURRENT FILING DATE: 1998-07-22
EARLIER APPLICATION NUMBER: PCT/US96/17081
EARLIER FILING DATE: 1996-10-22
EARLIER APPLICATION NUMBER: 08/589,711
EARLIER FILING DATE: 1996-01-22
EARLIER APPLICATION NUMBER: 08/427,023
EARLIER FILING DATE: 1995-04-24
NUMBER OF SEQ ID NOS: 66
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 24
LENGTH: 561
TYPE: DNA
ORGANISM: Borrelia burgdorferi
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(561)
US-09-117-257-24

Query Match 1.3%; Score 34.6; DB 4; Length 561;
Best Local Similarity 46.8%; Pred. No. 0.5;
Matches 109; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

QY 1765 GATTCACCTGCCAATGCTAGTACAGTTTATATCCCTAGGAAATTAATTACATTAGGGCT 1824
DB 299 GCTTCCTCTCTATGCTATAGCAATTTTCCGCTACTAGTACTGCACTTTTGGCT 240
QY 1825 GCACCGAACAATTGACAGCTGCTCAAAACCAATCTGTAGCGAGCTCTCAGAAATCT 1884
DB 239 TCAAGTATGATGATGATTTCTGATACCCCACTACCTTTTATCTTAAAGCATCA 180
QY 1885 AATAATGACAGCGCTCAAGCTGGGGTTAGTTAACTCCGAGGCTTTAGCCACTGCGCA 1944
DB 179 AATTTACACCCATAGAGAGCGCTTTTATATTCATCTATTCATCTACAAATGGCT 120
QY 1945 TCTATTCCTCCGCAATCTCTCAACCTGCTGCCCTGAGAGTACCAACTAT 1997
DB 119 TTAGCTGATGATTTCTATTTGATTTTGTGCTGCTGTTAGTCCACATGATAT 67

RESULT 12
US-09-221-938-3/c
Sequence 3, Application US/09221938
Patent No. 6228835

GENERAL INFORMATION:
APPLICANT: Guo, Betty
TITLE OF INVENTION: DECORIN BINDING PROTEIN COMPOSITIONS
TITLE OF INVENTION: AND METHODS OF USE
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: United States of America
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/221,938
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/589,711
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928

REFERENCE/DOCKET NUMBER: TAMK-158
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (512) 474-7577
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 561 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-221-938-3

Query Match 1.3%; Score 34.6; DB 4; Length 561;
Best Local Similarity 46.8%; Pred. No. 0.5;
Matches 109; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

QY 1765 GATTCACCTGCCAATGCTAGTACAGTTTATATCCCTAGGAAATTAATTACATTAGGGCT 1824
DB 299 GCTTCCTCTCTATGCTATAGCAATTTTCCGCTACTAGTACTGCACTTTTGGCT 240
QY 1825 GCACCGAACAATTGACAGCTGCTCAAAACCAATCTGTAGCGAGCTCTCAGAAATCT 1884
DB 239 TCAAGTATGATGATGATTTCTGATACCCCACTACCTTTTATCTTAAAGCATCA 180
QY 1885 AATAATGACAGCGCTCAAGCTGGGGTTAGTTAACTCCGAGCTTTAGCCACTGCGCA 1944
DB 179 AATTTACACCCATAGAGAGCGCTTTTATATTCATCTATTCATCTACAAATGGCT 120
QY 1945 TCTATTCCTCCGCAATCTCTCAACCTGCTGCCCTGAGAGTACCAACTAT 1997
DB 119 TTAGCTGATGATTTCTATTTGATTTTGTGCTGCTGTTAGTCCACATGATAT 67

RESULT 13
US-08-945-476-8/c
Sequence 8, Application US/08945476
Patent No. 6248517

GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: DECORIN BINDING PROTEIN COMPOSITIONS AND
TITLE OF INVENTION: METHODS OF USE
NUMBER OF SEQUENCES: 27
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/945,476
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/589,711
FILING DATE: 22-JAN-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/427,023
FILING DATE: 24-APR-1995
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 561 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-945-476-8

Query Match 1.3%; Score 34.6; DB 4; Length 561;
Best Local Similarity 46.8%; Pred. No. 0.5;
Matches 109; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

QY 1765 GATTCACCTGCCAATGCTAGTACAGTTTATATCCCTAGGAAATTAATTACATTAGGGCT 1824
DB 299 GCTTCCTCTCTATGCTATAGCAATTTTCCGCTACTAGTACTGCACTTTTGGCT 240
QY 1825 GCACCGAACAATTGACAGCTGCTCAAAACCAATCTGTAGCGAGCTCTCAGAAATCT 1884

```

Db      239 TCAAGTATGATGATTTTCTGATACCCAGTACCCGTTTATTAATTAAAGCATCA 180
      1885 AATAATGACGCGCCCTCAAGCTGGGGTGTAGTTAACTCCGAGCTTTTACCATCTGGCA 1944
      179 AAATTTACACCATAGAGAGAGCCCTTTTAAATTCATCTATTTCAATTCATGAGCT 120
Qy      1945 TCTATTTCCCTGCAACTTTCACACCTGCTGCGCTGAGAGTACCAACCTAT 1997
      119 TTACGTGATGATTTCTAATTTGATTTTGTGCTCTCTGTTAGTCACATGATAT 67

```

RESULT 14
US-08-945-476-22/c
Sequence 22, Application US/08945476
Patent No. 6248517

GENERAL INFORMATION:

APPLICANT:
TITLE OF INVENTION: DECORIN BINDING PROTEIN COMPOSITIONS AND
METHODS OF USE
NUMBER OF SEQUENCES: 27
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/945,476

FILING DATE:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/589,711
FILING DATE: 22-JAN-1996

APPLICATION NUMBER: US 08/427,023
FILING DATE: 24-APR-1995

INFORMATION FOR SEQ ID NO: 22:

SEQUENCE CHARACTERISTICS:
LENGTH: 561 base pairs

TYPE: nucleic acid
STRANDEDNESS: single

TOPOLOGY: linear

FEATURE:
NAME/KEY: CDS
LOCATION: 1..561

US-08-945-476-22

Query Match 1.3%; Score 34.6; DB 4; Length 561;
Best Local Similarity 46.8%; Pred. No. 0.5;

Matches 109; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

```

Qy      1765 GATTCACTGCCAATGCTAGTACAGTTATATCTCTAGGAAATTAATTAGGGCT 1824
      299 GCTTCCTCTTCTATTTGCTATTACGAAATTTTCCGCTACTGTAGTACGACCTTTTGT 240
Db      1825 GCACCAAGAACATTTGACAGCTGCTTCAAAACCATCTGTAGGAGCTCTCAGATACCT 1884
      239 TCAAGTATGATGATTTTGTATACCCACACCGCTTTTATCTTTAAAGCATCA 180
Qy      1885 AATAATGACGCGCTCAAGCTGGGGTGTAGTTAACTCCGAGCTTTTACCATCTGGCA 1944
      179 AAATTTACACCATAGAGAGAGCCCTTTTAAATTCATCTATTTCAATTCATGAGCT 120
Db      1945 TCTATTTCCCTGCAACTTTCACACCTGCTGCGCTGAGAGTACCAACCTAT 1997
      119 TTACGTGATGATTTCTAATTTGATTTTGTGCTCTCTGTTAGTCACATGATAT 67

```

RESULT 15
US-08-945-476-24/c
Sequence 24, Application US/08945476
Patent No. 6248517
GENERAL INFORMATION:
APPLICANT:

```

      TITLE OF INVENTION: DECORIN BINDING PROTEIN COMPOSITIONS AND
      METHODS OF USE
      NUMBER OF SEQUENCES: 27
      COMPUTER READABLE FORM:
      MEDIUM TYPE: Floppy disk
      COMPUTER: IBM PC compatible
      OPERATING SYSTEM: PC-DOS/MS-DOS
      SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
      CURRENT APPLICATION DATA:
      APPLICATION NUMBER: US/08/945,476
      FILING DATE:
      PRIOR APPLICATION DATA:
      APPLICATION NUMBER: US 08/589,711
      FILING DATE: 22-JAN-1996
      PRIOR APPLICATION DATA:
      APPLICATION NUMBER: US 08/427,023
      FILING DATE: 24-APR-1995
      INFORMATION FOR SEQ ID NO: 24:
      SEQUENCE CHARACTERISTICS:
      LENGTH: 561 base pairs
      TYPE: nucleic acid
      STRANDEDNESS: single
      TOPOLOGY: linear
      FEATURE:
      NAME/KEY: CDS
      LOCATION: 1..561
      US-08-945-476-24

```

Query Match

Best Local Similarity 1.3%; Score 34.6; DB 4; Length 561;
Pred. No. 0.5;

Matches 109; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

```

Qy      1765 GATTCACTGCCAATGCTAGTACAGTTATATCTCTAGGAAATTAATTAGGGCT 1824
      299 GCTTCCTCTTCTATTTGCTATTACGAAATTTTCCGCTACTGTAGTACGACCTTTTGT 240
Db      1825 GCACCAAGAACATTTGACAGCTGCTTCAAAACCATCTGTAGGAGCTCTCAGATACCT 1884
      239 TCAAGTATGATGATTTTGTATACCCACACCGCTTTTATCTTTAAAGCATCA 180
Qy      1885 AATAATGACGCGCTCAAGCTGGGGTGTAGTTAACTCCGAGCTTTTACCATCTGGCA 1944
      179 AAATTTACACCATAGAGAGAGCCCTTTTAAATTCATCTATTTCAATTCATGAGCT 120
Db      1945 TCTATTTCCCTGCAACTTTCACACCTGCTGCGCTGAGAGTACCAACCTAT 1997
      119 TTACGTGATGATTTCTAATTTGATTTTGTGCTCTCTGTTAGTCACATGATAT 67

```

Search completed: June 19, 2003, 04:27:58
Job time : 138 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 18, 2003, 23:11:38 : Search time 601 Seconds
(without alignments)
10139.613 Million cell updates/sec

Title: US-09-920-705-2

Perfect score: 2706
Sequence: 1 atggcgtactatgaagcc.....cgctgacaggggaccttga 2706

Scoring table:
IDENTITY-NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

N_Geneseq_101002:*

- 1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
- 2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
- 3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
- 4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
- 5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
- 6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
- 7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
- 8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
- 9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
- 10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
- 11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
- 12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
- 13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
- 14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
- 15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
- 16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
- 17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
- 18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
- 19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
- 20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
- 21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
- 22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
- 23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
- 24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	2706	100.0	2706	24	ABA95167 Arabidopsis floral
2	1896.8	70.1	4593	24	ABA95167 Arabidopsis floral
3	797.8	29.5	801	24	ABA95168 Arabidopsis floral
4	271.4	10.0	349	24	ABA95169 Arabidopsis floral
5	271.4	10.0	3715	24	ABA95170 Arabidopsis floral
6	113.2	4.2	438	24	ABL94207 Arabidopsis thalia
7	67.6	2.5	865	21	AAC53156 Arabidopsis thalia
8	47	1.7	4590	22	AAH24065 Yeast AOD9604-asso
9	42.8	1.6	256	23	AAS58705 CDNA #1381 encodin

10	42.2	1.6	1595	24	AB199579
11	42.2	1.6	1893	22	AAH14157
12	41.6	1.5	2142	24	ABK63660
13	41.4	1.5	2052	24	AB199466
14	41	1.5	11680	21	AAC76700
15	39.6	1.5	6104	24	ABL33124
16	39	1.4	2190	21	AAZ51263
17	38.4	1.4	4668	22	AAH98410
18	38.4	1.4	5134	23	ABV24463
19	38	1.4	6161	23	ABL09562
20	37.8	1.4	4561	22	AAF32424
21	37.8	1.4	4907	22	AAF32416
22	37.8	1.4	4946	11	AAO03115
23	37.8	1.4	5063	22	AAF32425
24	36.4	1.3	331	21	AAC01628
25	36.4	1.3	396	22	AAS37650
26	36.4	1.3	401	21	AAH30845
27	36.4	1.3	406	22	AAS37725
28	36.4	1.3	407	22	AAS37723
29	36.4	1.3	631	24	ABO56487
30	36.4	1.3	808	22	AAH05825
31	36.4	1.3	1501	22	AAH13847
32	36.4	1.3	3080	22	AAC91322
33	36.4	1.3	4069	24	ABN59810
34	36	1.3	4590	22	AAH24065
35	36	1.3	580073	18	AAT58840
36	35.8	1.3	6015	24	ABL32677
37	35.8	1.3	46340	21	ABN97978
38	35.6	1.3	1484	21	AAC47540
39	35.6	1.3	1486	21	AAC40456
40	35.6	1.3	2155561	24	ABN71557
41	35.4	1.3	3489	21	AAH30290
42	35.4	1.3	3489	22	AAH82901
43	35.4	1.3	3489	24	ABA93487
44	35.4	1.3	5059	20	AAH84332
45	35.4	1.3	32207	20	AAV73805

ALIGNMENTS

RESULT 1					
ID	ABA95167	standard:	CDNA: 2706 BP.		
XX	ABA95167:				
AC	20-MAY-2002	(first entry)			
XX					
DE	Arabidopsis floral induction gene (FPA) CDNA sequence.				
XX	FPA: floral induction; Photoperiod; plant; flowering; FLC; FRI; frigidia;				
KW	flowering locus C; transgenic; gene; ss.				
XX					
OS	Arabidopsis thaliana.				
XX					
EH	key	Location/Qualifiers			
FT	CDS	1..2706			
FT		/*tag= a			
FT	misc_binding	/product= "FPA protein"			
FT		46..279			
FT		/*tag= b			
FT	misc_binding	/note= "RNA binding region"			
FT		283..522			
FT		/*tag= c			
FT	misc_binding	/note= "RNA binding region"			
FT		610..852			
FT		/*tag= d			
FT		/note= "RNA binding region"			
PN	W0200212518-A2.				
XX					
PD	14-FEB-2002.				

Mouse ischaemic co
Human cDNA sequenc
Rat sequence diffe
Mouse ischaemic co
Human ORF ORF2255
Human immune syste
Human RNA-associat
Rat EST-derived co
Human prostate exp
Drosophila melanog
L. mesenteroides s
Nucleotide sequenc
L. mesenteroides s
Human secreted pro
Novel human diagno
Human colon cancer
Human colon cancer
Novel human diagno
Human human diagno
Human colon cancer
Human cDNA clone (H
Human cDNA sequenc
Human polynucleoti
Novel human coding
Yeast AOD9604-asso
Mycoplasma genital
Human immune syste
Human retroviral s
Arabidopsis thalia
Streptococcus poly
Kaposi's sarcoma-a
Nucleotide sequenc
Kaposi's sarcoma-a
Stealth virus nucl
KSHV LTR DNA (nucl

XX 02-AUG-2001; 2001WO-US24427.
 PF 03-AUG-2000; 2000US-22250P.
 PR (WISC) WISCONSIN ALUMNI RES FOUND.
 PA Amasino RM, Schomburg FM, Michaels SD, Patton D;
 PI WPI: 2002-227160/28.
 DR P-PSDB; ABB07659.
 XX
 PT Novel isolated DNA sequence comprising coding sequence for floral
 induction gene (FPA) gene which controls flowering time in plants,
 PT useful for altering flowering time of plants and for downregulating
 PT flowering locus C (FLC) mRNA activity
 XX
 PS Claim 5; Page 29-34; 39pp; English.
 CC The invention relates to the FPA gene (floral induction promoter in
 CC plants during both long and short day photoperiods) from Arabidopsis
 CC thaliana. The FPA coding sequence is useful for altering flowering time
 CC in a plant, and for down-regulating flowering locus C (FLC) mRNA
 CC activity. The FPA gene can accelerate flowering in several late-flowering
 CC mutant backgrounds and can fully compensate for addition of two naturally
 CC occurring genes, *frigida* (FRI) and *FLC* which confer late-flowering
 CC phenotypes. Over expression of FPA can compensate for the delaying effect
 CC caused by short days on floral induction, and decreases FLC mRNA in
 CC plants containing the FLC gene. Fragments of the FPA gene can also act to
 CC decrease activity of an endogenous FPA gene by modifying the expression
 CC of the endogenous FPA gene and expression of a portion of polypeptide
 CC encoded by FPA gene can also lead to a delay of flowering in a plant. The
 CC present sequence represents the *A. thaliana* FPA protein encoding cDNA.
 XX
 S0 Sequence 2706 BP; 779 A; 628 C; 617 G; 682 T; 0 other;
 Query Match 100.0%; Score 2706; DB 24; Length 2706;
 Best Local similarity 100.0%; Pred. No. 0;
 Matches 2706; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATGGCGTTATCTATGAAAGCCATTCAGAGCCGATGATTCGGTTTCCAGTCAACAATCTT 60
 DB 1 ATGGCGTTATCTATGAAAGCCATTCAGAGCCGATGATTCGGTTTCCAGTCAACAATCTT 60
 QY 61 TGGGTGGTACCTTAACGCCGAGAGACAGACAGATGATGACCGAGTTTGGAGA 120
 DB 61 TGGGTGGTACCTTAACGCCGAGAGACAGACAGATGATGACCGAGTTTGGAGA 120
 QY 121 TACGGCGATATGATAGATCAGGCTGATTTCTTCAAGAGGCTTTCGTTTATATCTAC 180
 DB 121 TACGGCGATATGATAGATCAGGCTGATTTCTTCAAGAGGCTTTCGTTTATATCTAC 180
 QY 181 AGACATGTGAGAGACAGATGCGCAAGAGAGGCTCTTCAAGAGCAATTTGAATGA 240
 DB 181 AGACATGTGAGAGACAGATGCGCAAGAGAGGCTCTTCAAGAGCAATTTGAATGA 240
 QY 241 AGCAATATTAAGTCAATATGAGACAGACCGGCAAACTTTGAAGTATATGAGTGA 300
 DB 241 AGCAATATTAAGTCAATATGAGACAGACCGGCAAACTTTGAAGTATATGAGTGA 300
 QY 301 GGAATCGGCCCTATATCTCAAGAGATGACCTGAGAGAGAGTTCGAAGTTTGGGAA 360
 DB 301 GGAATCGGCCCTATATCTCAAGAGATGACCTGAGAGAGAGTTCGAAGTTTGGGAA 360
 QY 361 ATCGAGAGATTTAGTCTTCTCAGAGACCAAGACAGCTTTCATGATATATATGAGATG 420
 DB 361 ATCGAGAGATTTAGTCTTCTCAGAGACCAAGACAGCTTTCATGATATATATGAGATG 420
 QY 421 GATGAGCTTTACAGGCTAAGAGCATGATGGAAGCCATATGGGTGATGCTTTTGGCT 480
 DB 421 GATGAGCTTTACAGGCTAAGAGCATGATGGAAGCCATATGGGTGATGCTTTTGGCT 480
 QY 481 GTTGATTTTCTCGGTACACAGCGCCAAAAAAGAACATATGGGTGCTCTTACATAC 540

DB 481 GTTGATTTTCTCGGTACACAGCGCCAAAAAAGAACATATGGGTGCTCTTACATAC 540
 QY 541 AGAATGCAATATGATCAATTAACCGCAGATCTCCTCATATGAGACTTAAAGA 600
 DB 541 AGAATGCAATATGATCAATTAACCGCAGATCTCCTCATATGAGACTTAAAGA 600
 QY 601 GATGTCAGCCCAAGTAAGTGTGTGATGGGTTCCTCTCTACTGCTACCAATCAAT 660
 DB 601 GATGTCAGCCCAAGTAAGTGTGTGATGGGTTCCTCTCTACTGCTACCAATCAAT 660
 QY 661 GATGAGCAATTTCTGCAATATGAGATATCTCTTGTGTAGATCGAGAGGTTAAAGT 720
 DB 661 GATGAGCAATTTCTGCAATATGAGATATCTCTTGTGTAGATCGAGAGGTTAAAGT 720
 QY 721 TACCATCAAGAAATTTTGCATCTTGAGATTGAGAGCGCGAGAGAGAGCTGCCAATG 780
 DB 721 TACCATCAAGAAATTTTGCATCTTGAGATTGAGAGCGCGAGAGAGAGCTGCCAATG 780
 QY 781 AAGGAGGCTTACAGGGGAGGTTATTCATATATCTAGAAATCAAAATTAATGACTCAAC 840
 DB 781 AAGGAGGCTTACAGGGGAGGTTATTCATATATCTAGAAATCAAAATTAATGACTCAAC 840
 QY 841 GATGAGTGCCTCTGAGCAAGACGATAGTATTTTACTGCTATGAAAGGTCACAG 900
 DB 841 GATGAGTGCCTCTGAGCAAGACGATAGTATTTTACTGCTATGAAAGGTCACAG 900
 QY 901 ACAGATATGTTCAATATATGATCTTCATGATCTTCTCTCATATCTAGAAATCTCT 960
 DB 901 ACAGATATGTTCAATATATGATCTTCATGATCTTCTCTCATATCTAGAAATCTCT 960
 QY 961 GGGTCTATGAGGCGCCCTCAGAGGTAGCAATGAGGCTTCATTAATGAGTGCACAAAT 1020
 DB 961 GGGTCTATGAGGCGCCCTCAGAGGTAGCAATGAGGCTTCATTAATGAGTGCACAAAT 1020
 QY 1021 GACGTTGTTGTAAGAGACCAAACTGAGAGAGCCATCGCAATATGAACTGGAATATCTC 1080
 DB 1021 GACGTTGTTGTAAGAGACCAAACTGAGAGAGCCATCGCAATATGAACTGGAATATCTC 1080
 QY 1081 CCATCTCCAAACAGAGACTGGAATCTCCCATCTCTCTGCAAGAGTACAGAGCGCCCTATG 1140
 DB 1081 CCATCTCCAAACAGAGACTGGAATCTCCCATCTCTCTGCAAGAGTACAGAGCGCCCTATG 1140
 QY 1141 AGGTCAAAACCCGATCTCTGGAGAGATATGATCTGCTCAGTGTGTAGAGAAATGTA 1200
 DB 1141 AGGTCAAAACCCGATCTCTGGAGAGATATGATCTGCTCAGTGTGTAGAGAAATGTA 1200
 QY 1201 CGAACCAGAAAGATGATGATCAGTGAACGGTTTACTCCAAATGAGTGTGATGAGAGTCA 1260
 DB 1201 CGAACCAGAAAGATGATGATCAGTGAACGGTTTACTCCAAATGAGTGTGATGAGAGTCA 1260
 QY 1261 TTTGTCGAGGTTCAAGTGTCTAGACCTATCCGTGCCCCCGCTGATTTGATCACA 1320
 DB 1261 TTTGTCGAGGTTCAAGTGTCTAGACCTATCCGTGCCCCCGCTGATTTGATCACA 1320
 QY 1321 TGGAGAGAAATGATCCCAAGAGGTGGAATCCCGTCTTGTGCTGCTGTTGTA 1380
 DB 1321 TGGAGAGAAATGATGATCCCAAGAGGTGGAATCCCGTCTTGTGCTGCTGTTGTA 1380
 QY 1381 GGAAGAGGATTTGAACCTTAACCTGCTGAGTCTGCAATTTCTACGCAAGAGTATG 1440
 DB 1381 GGAAGAGGATTTGAACCTTAACCTGCTGAGTCTGCAATTTCTACGCAAGAGTATG 1440
 QY 1441 AATATGCTGCTAAACATTTACGCGCTTCCATTTGAGATGAGATGATTTTGTGCA 1500
 DB 1441 AATATGCTGCTAAACATTTACGCGCTTCCATTTGAGATGAGATGATTTTGTGCA 1500
 QY 1501 GACAGGAGAAAGATTTTGGCTTACACTGAATTTCTCCGGTACTTAGCTCAAAAGAT 1560
 DB 1501 GACAGGAGAAAGATTTTGGCTTACACTGAATTTCTCCGGTACTTAGCTCAAAAGAT 1560
 QY 1561 CGGCGGGTGTTCGCAATTTGATGATGATGATCAACTTATTTCTGTGCTCCATCAGAT 1620

Db 2346 ATACTCTTGGATGAGATGAGAGGGTAAAGTTTCCATCAGAAATTTTGACCTTGTG 2405
Qy 748 GAGTTAGAGCGCGAGAAAGCTGCGCAATGCAAGAAAGGCTTACAGGGAGGTTATTC 807
Db 2406 GAGTTAGAGCGCGAGAAAGCTGCGCAATGCAAGAAAGGCTTACAGGGAGGTTATTC 2465
Qy 808 AATAATCTAGAAATTAATGATGCAACAGATGTTGCTCCCTGACACAGAGAT 867
Db 2466 AATAATCTAGAAATTAATGATGCAACAGATGTTGCTCCCTGACACAGAGAT 2525
Qy 868 ACTAGTTTACTCTGTATGAAACGGTCAAGAGACAGATATGTTCAATATGATCTTCA 927
Db 2526 ACTAGTTTACTCTGTATGAAACGGTCAAGAGACAGATATGTTCAATATGATCTTCA 2585
Qy 928 TGTATGCTGCTCCATCTTACAGAAATCTGAGAAATCTGGGCTATGAGAGGCTTACAGAGTACG 987
Db 2586 TGTATGCTGCTCCATCTTACAGAAATCTGAGAAATCTGGGCTATGAGAGGCTTACAGAGTACG 2645
Qy 988 AATGAGCTGTCATATATGATGTCAGAAATACAAATGACGTTGTTGTAAGAGCCAAAGTGG 1047
Db 2646 AATGAGCTGTCATATATGATGTCAGAAATACAAATGACGTTGTTGTAAGAGCCAAAGTGG 2705
Qy 1048 AGGAGGCAATCTGCAAAATGGAACCTGGAATCTCCATCTCCACAGAGACCTGGAATCTC 1107
Db 2706 AGGAGGCAATCTGCAAAATGGAACCTGGAATCTCCATCTCCACAGAGACCTGGAATCTC 2765
Qy 1108 CCATCTCTGTCAGAGGTAGAGAGGCGCTATGAGTCAAAACCCGATCTTGTGGAGGA 1167
Db 2766 CCATCTCTGTCAGAGGTAGAGAGGCGCTATGAGTCAAAACCCGATCTTGTGGAGGA 2825
Qy 1168 TATATCTCTGCTGCTGATGTCAGAAAGTAACGAACGAGAGATGATCAGTGGAC 1227
Db 2826 TATATCTCTGCTGCTGATGTCAGAAAGTAACGAACGAGAGATGATCAGTGGAC 2885
Qy 1228 GGTATCTCTCAATGGGTGTCGATGAGAGGTCATTTGAGGTTGAGTTCAGTTCGCTAGA 1287
Db 2886 GGTATCTCTCAATGGGTGTCGATGAGAGGTCATTTGAGGTTGAGTTCAGTTCGCTAGA 2945
Qy 1288 CCTATCCGTGGCCCCCTGATTTCTGATCACAATGAGAGAAATGATTCGAAGGTTGA 1347
Db 2946 CCTATCCGTGGCCCCCTGATTTCTGATCACAATGAGAGAAATGATTCGAAGGTTGA 3005
Qy 1348 ACTCCGCTGCTGCTGCTGCTGATACCATATGAGAAAGGAGTTGAATAACT--- 1403
Db 3006 ACTCCGCTGCTGCTGCTGCTGATACCATATGAGAAAGGAGTTGAATAACT--- 3065
Qy 1404 ----- 1403
Db 3066 GTACTAATTTCTAGCACTTTAACCCTTCTAGTCTTTCTTTTCAAGAGGATTTATAT 3125
Qy 1404 ----- 1403
Db 3126 TTTCCATTTTCATCTCGATGGAAGTAACATTTATATAGATGATTTTATTTACTA 3185
Qy 1404 ----- 1403
Db 3186 TTACTTTTGTAGTTCTGAGATGCTTGATTTTCAATGCTGTTGATTCATTTTGGCATTG 3245
Qy 1404 ----- 1403
Db 3246 CCCTCAATACTGACTTTGTTTTTTTTTAAATATGATTTATAGCCCTGAGGTGCTCA 3305
Qy 1419 TTGTTAGCAAGAACTGATTTGAATGCTGCTAAACATTAACGCGCTTGCCATTGAGATG 1478
Db 3306 TTGTTAGCAAGAACTGATTTGAATGCTGCTAAACATTAACGCGCTTGCCATTGAGATG 3365
Qy 1479 TGAGATCGTTTTTTCGTACCAAGAGAGAAAGATTTGCGTCTTACACTGAATTTCT 1538
Db 3366 TGAGATCGTTTTTTCGTACCAAGAGAGAAAGATTTGCGTCTTACACTGAATTTCT 3425
Qy 1539 CCGGTACTTACCTCAAAAGATCGGCGGCTTCCCAATATGATGATGATCAACTTT 1598

Db 3426 CCGGTACTTACCTCAAAAGATCGGCGGCTTCCCAATATGATGATGATCAACTTT 3485
Qy 1599 ATTCTGTGCTCCTCATCAGATTTCTTAACGATGATCTCCAGATGACCGCTCAAGAC 1658
Db 3486 ATTCTGTGCTCCTCATCAGATTTCTTAACGATGATCTCCAGATGACCGCTCAAGAC 3545
Qy 1659 GCTATATGTTGTTCTTCAAGTTACCCCGGAGCGCTTCTGTTACAGCATATACAG 1718
Db 3546 GCTATATGTTGTTCTTCAAGTTACCCCGGAGCGCTTCTGTTACAGCATATACAG 3605
Qy 1719 ACAAGATCTCAGTCCATCTCCTCGATTTATGATCAAGCCCGGATTCACCTGCCA 1778
Db 3606 ACAAGATCTCAGTCCATCTCCTCGATTTATGATCAAGCCCGGATTCACCTGCCA 3665
Qy 1779 TGTAGTACAGTTTATATCTCCTAGGAAATTAATGATGAGGCTGACACGAACATTT 1838
Db 3666 TGTAGTACAGTTTATATCTCCTAGGAAATTAATGATGAGGCTGACACGAACATTT 3725
Qy 1839 GACAGCTGCTTCAAAACCATCTGTATGCGAGCCTTCAGAAATCTAATATGACAGGCC 1898
Db 3726 GACAGCTGCTTCAAAACCATCTGTATGCGAGCCTTCAGAAATCTAATATGACAGGCC 3785
Qy 1899 TCAAGCTGGGTTAGTTTAACTCCGAGGCTTTAGCCAGTGGAGATCTATCTCCCTG 1958
Db 3786 TCAAGCTGGGTTAGTTTAACTCCGAGGCTTTAGCCAGTGGAGATCTATCTCCCTG 3845
Qy 1959 AACTTTCACCTGCTGCTGCTGAGAGTACCAACCTATGTCAGACCTTCAACAGTTGT 2018
Db 3846 AACTTTCACCTGCTGCTGCTGAGAGTACCAACCTATGTCAGACCTTCAACAGTTGT 3905
Qy 2019 TTCCACAGACATCACTCCCAATGAGTATGCAATGAGAGAAACCGTCCAGCTTGGAA 2078
Db 3906 TTCCACAGACATCACTCCCAATGAGTATGCAATGAGAGAAACCGTCCAGCTTGGAA 3965
Qy 2079 AAGAGTCCACAAACAGTTATGATGCTGCAATACGATCTTCAACATACGAAATCA 2138
Db 3966 AAGAGTCCACAAACAGTTATGATGCTGCAATACGATCTTCAACATACGAAATCA 4025
Qy 2139 GTACACTCCAGCTGGGCACTACCTCCTCTGCTGCTGCTTACCTCCAGCTTCAACAA 2198
Db 4026 GTACACTCCAGCTGGGCACTACCTCCTCTGCTGCTGCTTACCTCCAGCTTCAACAA 4085
Qy 2199 CCCCAACTACATAGTGAATGTCATAGGCAATGCAATACGAGGCAATCTGTTAA 2258
Db 4086 CCCCAACTACATAGTGAATGTCATAGGCAATGCAATACGAGGCAATCTGTTAA 4145
Qy 2259 CATGCTCAGCTGCTCCTTACCAATATGCTCAATATATATTTTCCATGTCATCA 2318
Db 4146 CATGCTCAGCTGCTCCTTACCAATATGCTCAATATATATTTTCCATGTCATCA 4205
Qy 2319 GGGTTCGCAATCAATCTGTTCTCAGCCCATGTCAGCAATACCAAGCAAGAGCTC 2378
Db 4206 GGGTTCGCAATCAATCTGTTCTCAGCCCATGTCAGCAATACCAAGCAAGAGCTC 4265
Qy 2379 CATGCCAAACCAAACTATGTCATTTCCAAATGATCAGCAAGCTATTTTCAATGGGCT 2438
Db 4266 CATGCCAAACCAAACTATGTCATTTCCAAATGATCAGCAAGCTATTTTCAATGGGCT 4325
Qy 2439 AACACCAATCGAGGACAGAACTTAACCTTCCCAATTTCAACTGTCATGCAACACC 2498
Db 4326 AACACCAATCGAGGACAGAACTTAACCTTCCCAATTTCAACTGTCATGCAACACC 4385
Qy 2499 AGCAGATTAAGCAAAATTTAGAGCCCAAAACCAAGCACTAGATTTGAGCTATGATCTC 2558
Db 4386 AGCAGATTAAGCAAAATTTAGAGCCCAAAACCAAGCACTAGATTTGAGCTATGATCTC 4445
Qy 2559 TGGGATGCTGAGGTTACAAAGATGGGAGGTGATTAAGATAGAGATACAGTCAAC 2618
Db 4446 TGGGATGCTGAGGTTACAAAGATGGGAGGTGATTAAGATAGAGATACAGTCAAC 4505
Qy 2619 ACTACAAATTTGACAGAAACCTTCTTCCCAATACAGAGAAAGAGAGCAAGAGCTTTC 2678
Db 4506 ACTACAAATTTGACAGAAACCTTCTTCCCAATACAGAGAAAGAGAGCAAGAGCTTTC 4565

0Y	2679	AGTACTCCGGCTGGACAGGGGCTTGA	2706
Db	4566	AGTACTCCGGCTGGACAGGGGCTTGA	4593
RESULT 3			
ID	ABA95168/C		
XX	ABA95168 standard; DNA; 801 BP.		
XX	ABA95168;		
XX	20-MAY-2002 (first entry)		
XX	Arabidopsis floral induction gene (FPA) antisense fragment;		
DE	FPA; floral induction; photoperiod; plant; flowering; FLC; FRI; frigida;		
XX	flowering locus C; transgenic; gene; antisense; ds.		
XX	Arabidopsis thaliana.		
XX	WO200212518-A2.		
XX	14-FEB-2002.		
XX	02-AUG-2001; 2001WO-US24427.		
XX	03-AUG-2000; 2000US-222550P.		
XX	(WISC) WISCONSIN ALUMNI RES FOUND.		
XX	Amasino RM, Schomburg FM, Michaels SD, Patton D;		
PI	WPI: 2002-227160/28.		
XX			
XX	Novel isolated DNA sequence comprising coding sequence for floral		
PT	induction gene (FPA) gene which controls flowering time in plants,		
PT	useful for altering flowering time of plants and for downregulating		
PR	flowering locus C (FLC) mRNA activity		
XX			
XX	Claim 5; Page 37; 39pp; English.		
XX			
CC	The invention relates to the FPA gene (floral induction promoter in		
CC	plants during both long and short day photoperiods) from Arabidopsis		
CC	thaliana. The FPA coding sequence is useful for altering flowering time		
CC	in a plant, and for down-regulating flowering locus C (FLC) mRNA		
CC	activity. The FPA gene can accelerate flowering in several late-flowering		
CC	mutant backgrounds and can fully compensate for addition of two naturally		
CC	occurring genes, frigida (FRI) and FLC which confer late-flowering		
CC	phenotypes. Over expression of FPA can compensate for the delaying effect		
CC	caused by short days on floral induction, and decreases FLC mRNA in		
CC	plants containing the FLC gene. Fragments of the FPA gene can also act to		
CC	decrease activity of an endogenous FPA gene by modifying the expression		
CC	of the endogenous FPA gene and expression of a portion of polypeptide		
CC	encoded by FPA gene can also lead to a delay of flowering in a plant. The		
CC	present sequence represents the A. thaliana FPA gene antisense fragment.		
XX			
XX	Sequence 801 BP; 206 A; 207 C; 166 G; 222 T; 0 other;		
QY			
	Query Match	29.5%; Score 797.8; DB 24; Length 801;	
	Best Local Similarity	99.8%; Pred. No. 1.6e-240;	
	Matches 799; Conservative 0; Mismatches 2; Indels 0; Gaps 0;		
Db			
	587	AGAGCTTAAAGGAGATGCCAGCAATAAGTCTGATGGATGGGTTCCCTCCTACTG	646
	801	AAGACTTTAAAGGAGATGCCAGCAATAAGTCTGATGGATGGGTTCCCTCCTACTG	742
QY			
	647	CTACACAAATGCATGATGAGCAAAATTCGCACAAATGCAGATGATACTTTGGTGAGATCG	706
Db			
	741	CTACACAAATGCATGATGAGCAAAATTCGCACAAATGCATGATGATCTTTGGTGAGATCG	682
QY			
	707	AGAGGGTAAAGATTAACCATCAAGGAATTTGGCACTTGTGAGATTAAAGAGCGCGGAGG	766

Db	681	AGAGGGTAAAAAGTTACCCATCAAGAAATTTTGCACCTTGTTGGAGTTTAAGAGCCGGAGC	622
QY	767	AAGCTCGCCATGCAAGGAAGGCCCTACAGGGGAGGTTATTCATATATCTAGATCAAAA	826
Db	621	AAGCTCGCCATGCAAGGAAGGCCCTACAGGGGAGGTTATTCATATATCTAGATCAAAA	562
QY	827	TTATGTACTCAACAGTAGTGGCTCTCTGAGCAAGACAGATCTAGTTTTTACTCTGTGA	886
Db	561	TTATGTACTCAACAGTAGTGGCTCTCTGAGCAAGACAGATCTAGTTTTTACTCTGTGA	502
QY	887	TGAACCGGTCAAGGACAGATGTGTTCATATATGATTCCTCATGTGTATCTCTCCATATG	946
QY	947	CTACTGCAATTCCTGGGTCTATGTAGAGGCCCTCAAGAGGTACGCAATGAGCGTTTCATATATG	1006
Db	441	CTACTGCAATTCCTGGGTCTATGTAGAGGCCCTCAAGAGGTACGCAATGAGCGTTTCATATATG	382
QY	1007	GTGCAAGATACAAATGACGTTGTTGGTAAAGAGCCAAACTGGAGGAGGCCATCTGCAAAATG	1066
Db	381	GTGCAAGATACAAATGACGTTGTTGGTAAAGAGCCAAACTGGAGGAGGCCATCTGCAAAATG	322
QY	1067	GAACTGCAATCTCTCCATCTCCACAGAGACCTGGAAATCTCCATCTCCGTGACACAGGTA	1126
Db	321	GAACTGCAATCTCTCCATCTCCACAGAGACCTGGAAATCTCCATCTCCGTGACACAGGTA	262
QY	1127	CGAGGCGCCCTATGAGTCAAAACCCCGATTTCTTGGGAAGATATGATCTCTCAGTTGG	1186
Db	261	CGAGGCGCCCTATGAGTCAAAACCCCGATTTCTTGGGAAGATATGATCTCTCAGTTGG	202
QY	1187	TCAGAGAAAGTAAACGAACAGAAAGAGATGATCAGTGGACGGTTTTACTCCAATGGGTG	1246
Db	201	TCAGAGAAAGTAAACGAACAGAAAGAGATGATCAGTGGACGGTTTTACTCCAATGGGTG	142
QY	1247	TCGATGAGAGGTCAATTGGTGTGAGGTTCAAGTGTGCTAGACCTATCCGTGGCCCCCTG	1306
Db	141	TCGATGAGAGGTCAATTGGTGTGAGGTTCAAGTGTGCTAGACCTATCCGTGGCCCCCTG	82
QY	1307	ATTTCGATCACAATATGAGAGGAATGATGTCGAAGGGTGAACCTCCGTCGTGTGCTC	1366
Db	81	ATTTCGATCACAATATGAGAGGAATGATGTCGAAGGGTGAACCTCCGTCGTGTGCTC	22
QY	1367	GTTGTGTACCTATGGGAAAG	1387
Db	21	GTTGTGTACCTATGGGAAAG	1
RESULT 4			
ABA95169			
ID	ABA95169	standard; DNA; 349 BP.	
XX	ABA95169;		
AC			
XX	20-MAY-2002 (first entry)		
XX			
DE	Arabidopsis floral induction gene (FPA) coding fragment.		
XX			
KM	FPA, floral induction; photoperiod; plant; flowering; FLC; FRI; friglida;		
KM	flowering locus C; transgenic; gene; ds.		
XX			
OS	Arabidopsis thaliana.		
XX			
PN	WO200212518-A2.		
PD			
XX	14-FEB-2002.		
XX			
PF	02-AUG-2001; 2001WO-US24427.		
XX			
PR	03-AUG-2000; 2000US-22250P.		
XX			
PA	(WISC) WISCONSIN ALUMNI RES FOUND.		
XX			
PI	Amesino RM, Schomburg FM, Michaels SD, Patton D;		

XX WPI: 2002-227160/28.

XX Novel isolated DNA sequence comprising coding sequence for floral
PT induction gene (FPA) gene which controls flowering time in plants,
PT useful for altering flowering time of plants and for downregulating
XX flowering locus C (FLC) mRNA activity

PS Claim 5; Page 37; 39pp; English.

XX The invention relates to the FPA gene (floral induction promoter in
CC plants during both long and short day photoperiods) from Arabidopsis
CC thaliana. The FPA coding sequence is useful for altering flowering time
CC in a plant, and for down-regulating flowering locus C (FLC) mRNA
CC activity. The FPA gene can accelerate flowering in several late-flowering
CC mutant backgrounds and can fully compensate for addition of two naturally
CC occurring genes, *frigida* (FRI) and *FLC* which confer late-flowering
CC phenotypes. Over expression of FPA can compensate for the delaying effect
CC caused by short days on floral induction, and decreases FLC mRNA in
CC plants containing the FLC gene. Fragments of the FPA gene can also act to
CC decrease activity of an endogenous FPA gene by modifying the expression
CC of the endogenous FPA gene and expression of a portion of polypeptide
CC encoded by FPA gene can also lead to a delay of flowering in a plant. The
CC present sequence represents the A. thaliana FPA gene coding fragment.

XX Sequence 349 BP; 92 A; 68 C; 81 G; 108 T; 0 other;

Query Match 10.0%; Score 271.4; DB 24; Length 349;

Best Local Similarity 94.6%; Pred. No. 1.1e-74;

Matches 281; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 1 ATGGCGTTATCTATGAAAGCCATTCAGAGCCGATGATTCGGTTCCAGTCAACAATCTT 60
DB 1 ATGGCGTTATCTATGAAAGCCATTCAGAGCCGATGATTCGGTTCCAGTCAACAATCTT 60
QY 61 TGGGTGGTAGCCCTTAAGCGCGGAGACAGACAGATCTGACCGAGTTGTTGAAGA 120
DB 61 TGGGTGGTAGCCCTTAAGCGCGGAGACAGACAGATCTGACCGAGTTGTTGAAGA 120
QY 121 TACGGCGATATGATAGCAATCAAGGCTGATCTTCAAGAGCTTTGGTTATATAC 180
DB 121 TACGGCGATATGATAGCAATCAAGGCTGATCTTCAAGAGCTTTGGTTATATAC 180
QY 181 AGCAGTGTGAG 240
DB 181 AGCAGTGTGAG 240
QY 241 AGTCAATTAAGATGCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 297
DB 241 AGTCAATTAAGATGCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 297

RESULT 5

ABA95170 standard; DNA; 3715 BP.

AC ABA95170;

DT 20-MAY-2002 (first entry)

DE Arabidopsis floral induction gene (FPA) fragment.

KW FPA; floral induction; photoperiod; plant; flowering; FLC; FRI; *frigida*;
KM flowering locus C; transgenic; gene; ds.

OS Arabidopsis thaliana.

PN MO200212518-A2.

XX 14-FEB-2002.

PF 02-AUG-2001; 2001WO-US24427.

XX

PR 03-AUG-2000; 2000US-222550P.

XX (WISC) WISCONSIN ALUMNI RES FOUND.

XX Anasino RM, Schomburg FM, Michaels SD, Patton D;

DR WPI: 2002-227160/28.

XX Novel isolated DNA sequence comprising coding sequence for floral
PT induction gene (FPA) gene which controls flowering time in plants,
PT useful for altering flowering time of plants and for downregulating
XX flowering locus C (FLC) mRNA activity

PS Claim 5; Page 38; 39pp; English.

XX The invention relates to the FPA gene (floral induction promoter in
CC plants during both long and short day photoperiods) from Arabidopsis
CC thaliana. The FPA coding sequence is useful for altering flowering time
CC in a plant, and for down-regulating flowering locus C (FLC) mRNA
CC activity. The FPA gene can accelerate flowering in several late-flowering
CC mutant backgrounds and can fully compensate for addition of two naturally
CC occurring genes, *frigida* (FRI) and *FLC* which confer late-flowering
CC phenotypes. Over expression of FPA can compensate for the delaying effect
CC caused by short days on floral induction, and decreases FLC mRNA in
CC plants containing the FLC gene. Fragments of the FPA gene can also act to
CC decrease activity of an endogenous FPA gene by modifying the expression
CC of the endogenous FPA gene and expression of a portion of polypeptide
CC encoded by FPA gene can also lead to a delay of flowering in a plant. The
CC present sequence represents the A. thaliana FPA gene fragment comprising
XX the FPA promoter and intron.

XX Sequence 3715 BP; 1055 A; 728 C; 711 G; 1221 T; 0 other;

Query Match 10.0%; Score 271.4; DB 24; Length 3715;

Best Local Similarity 94.6%; Pred. No. 4.9e-74;

Matches 281; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 1 ATGGCGTTATCTATGAAAGCCATTCAGAGCCGATGATTCGGTTCCAGTCAACAATCTT 60
DB 1832 ATGGCGTTATCTATGAAAGCCATTCAGAGCCGATGATTCGGTTCCAGTCAACAATCTT 1891
QY 61 TGGGTGGTAGCCCTTAAGCGCGGAGACAGACAGATCTGACCGAGTTGTTGAAGA 120
DB 1892 TGGGTGGTAGCCCTTAAGCGCGGAGACAGACAGATCTGACCGAGTTGTTGAAGA 1951
QY 121 TACGGCGATATGATAGCAATCAAGGCTGATCTTCAAGAGCTTTGGTTATATAC 180
DB 1952 TACGGCGATATGATAGCAATCAAGGCTGATCTTCAAGAGCTTTGGTTATATAC 2011
QY 181 AGCAGTGTGAG 240
DB 2012 AGCAGTGTGAG 2071
QY 241 AGTCAATTAAGATGCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 297
DB 2072 AGTCAATTAAGATGCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2128

RESULT 6

ABL94207 standard; cDNA; 438 BP.

AC ABL94207;

DT 10-JUN-2002 (first entry)

DE Arabidopsis thaliana nucleic acid sequence Ref:2027972 SEQ ID NO:972.

KW Arabidopsis thaliana; insecticide; fungicide; plant; mapping; diagnosis;

XX genetic modification; gene; ss.

OS Arabidopsis thaliana.

XX

PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139763.
PR 22-JUN-1999; 99US-0139817.
PR 23-JUN-1999; 99US-0139859.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140981.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145313.
PR 27-JUL-1999; 99US-0145318.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.

PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147182.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 25-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

[illegible]

RESULT 8
AAH24065
ID AAH24065 standard; DNA; 4590 BP

DT 29-AUG-2001 (first entry)

Yeast AOD9604-associated DNA sequence, SEQ ID NO:1

KW Human growth hormone analogue peptide; hGH; AOD6604; lipid metabolism;
KW modulation; lipolysis stimulation; hormone-sensitive lipase stimulation
KW lipogenesis inhibition; acetyl CoA carboxylase inhibition; obesity;
KW functional food; transgenic yeast; fat/lean ratio; food use; ds.

OS. *Saccharomyces cerevisiae*.

	key	Location/Qualifiers
FH	misc_feature	10
FT		

FT	/note-	"Represented as *	In the specification"
FT	misc_feature	3617	

	/note=	"Represented as *	In the specification"
FT	misc_feature		
FT			
FT			

FT	2	Represented as *	In the specification"
FT	3679		
misc_feature			

	misc_feature	FT	FT
	3819	"/note= "Represented as *	in the specification"

	FT	FT	FT
misc-feature	3862	3862	3862

	Represented as *	In the specification
FT	/note-	
FT	misc_feature	3864

	misc	feature
FT	3888	3888

2	FT	misc	feature	3890	Represented as * in the specification"
2	FT	misc	feature	3890	Represented as * in the specification"

FT	misc feature	FT
3912	"/note="	"Represented as * in the specification"

FT	misc	feature	3914	Represented as * in the specification
----	------	---------	------	---------------------------------------

FT	lag=	A	note=	* in the specification
1	1	0.000	0.000	0.000
2	2	0.000	0.000	0.000
3	3	0.000	0.000	0.000
4	4	0.000	0.000	0.000
5	5	0.000	0.000	0.000
6	6	0.000	0.000	0.000
7	7	0.000	0.000	0.000
8	8	0.000	0.000	0.000
9	9	0.000	0.000	0.000
10	10	0.000	0.000	0.000
11	11	0.000	0.000	0.000
12	12	0.000	0.000	0.000
13	13	0.000	0.000	0.000
14	14	0.000	0.000	0.000
15	15	0.000	0.000	0.000
16	16	0.000	0.000	0.000
17	17	0.000	0.000	0.000
18	18	0.000	0.000	0.000
19	19	0.000	0.000	0.000
20	20	0.000	0.000	0.000
21	21	0.000	0.000	0.000
22	22	0.000	0.000	0.000
23	23	0.000	0.000	0.000
24	24	0.000	0.000	0.000
25	25	0.000	0.000	0.000
26	26	0.000	0.000	0.000
27	27	0.000	0.000	0.000
28	28	0.000	0.000	0.000
29	29	0.000	0.000	0.000
30	30	0.000	0.000	0.000
31	31	0.000	0.000	0.000
32	32	0.000	0.000	0.000
33	33	0.000	0.000	0.000
34	34	0.000	0.000	0.000
35	35	0.000	0.000	0.000
36	36	0.000	0.000	0.000
37	37	0.000	0.000	0.000
38	38	0.000	0.000	0.000
39	39	0.000	0.000	0.000
40	40	0.000	0.000	0.000
41	41	0.000	0.000	0.000
42	42	0.000	0.000	0.000
43	43	0.000	0.000	0.000
44	44	0.000	0.000	0.000
45	45	0.000	0.000	0.000
46	46	0.000	0.000	0.000
47	47	0.000	0.000	0.000
48	48	0.000	0.000	0.000
49	49	0.000	0.000	0.000
50	50	0.000	0.000	0.000
51	51	0.000	0.000	0.000
52	52	0.000	0.000	0.000
53	53	0.000	0.000	0.000
54	54	0.000	0.000	0.000
55	55	0.000	0.000	0.000
56	56	0.000	0.000	0.000
57	57	0.000	0.000	0.000
58	58	0.000	0.000	0.000
59	59	0.000	0.000	0.000
60	60	0.000	0.000	0.000
61	61	0.000	0.000	0.000
62	62	0.000	0.000	0.000
63	63	0.000	0.000	0.000
64	64	0.000	0.000	0.000
65	65	0.000	0.000	0.000
66	66	0.000	0.000	0.000
67	67	0.000	0.000	0.000
68	68	0.000	0.000	0.000
69	69	0.000	0.000	0.000
70	70	0.000	0.000	0.000
71	71	0.000	0.000	0.000
72	72	0.000	0.000	0.000
73	73			

FT	misc_feature	3938	/tag=	1	"Represented as * in the specification"
FT			/note=		
FT					
FT	misc_feature	3939	/tag=	m	"Represented as * in the specification"
FT			/note=		
FT					
FT	misc_feature	3941	/tag=	o	"Represented as * in the specification"
FT			/note=		
FT					
FT	misc_feature	3943	/tag=	p	"Represented as * in the specification"
FT			/note=		
FT					
FT	misc_feature	4361	/tag=	q	"Represented as * in the specification"
FT			/note=		
FT					

PN WO200133977-A1

PD 17-MAY-2001.

PF 06-NOV-2000; 2000WO-AU01362.

PR 05-NOV-1999; 99AU-0003875.

PA (META-) METABOLIC PHARM LTD.

PI Belyea CI, Ng FM, Vaughan P;

DR WPI; 2001-328876/34.

PT New organisms containing nucleic acid encoding a growth hormone

PT aids for obesity and in the meat production industry

PS Disclosure; Page 48-50; 54pp; English

The invention relates to novel transgenic organisms useful in the production of functional food and drink products for the treatment or prevention of obesity via the regulation of lipid metabolism. The organisms comprise a polynucleotide encoding a growth hormone fragment capable of stimulating the activity of hormone-sensitive lipase (the key enzyme in lipolysis) and inhibiting acetyl CoA carboxylase (the key enzyme in lipogenesis). The growth hormone fragment preferably contains at least the disulphide-bonded loop of a mammalian growth hormone (but is not the full-length growth hormone) and is optionally linked to an epitope tag or heterologous fusion protein partner. The transgenic organism may be a microorganism used to produce a fermented product (e.g., yeast), or an edible plant or animal or cell thereof. Food or drink made using methods of the invention are used to modify fat/lean ratio, lipid metabolism or food use in a mammal. In particular, the food or drink products may be used to treat or prevent obesity, particularly in humans, and may also be used to improve the fat/lean ration of livestock raised for meat production. In the exemplification of the invention, the human growth hormone (hGH) fragment analogue AOD9604 was expressed in yeast, optionally fused to the FLAG epitope (AAB73625). The present sequence is described as a DNA sequence from yeast in the sequence listing, but is not further referred to in the specification.

Sequence 4590 BP; 661 A; 384 C; 127 G; 522 T; 2896 other;

Query Match 1.78; Score 47; DB 22; Length 4590;

Matches 100; Conservative 346; Mismatches 313; Indels 13; Gaps 3;

QY 116 GAAGATACGGCGATATTGATAGATCAGGCGTGTATTCTTCACGAGGCTTGGCTTATAT 175

Db 3586 SANDDRAWNGSDATHTSNTRDATDAYGSTNMRGMATTHDAYNTRMTHANDYARGC 3645

QY 176 ACTACAGACATGTGGAGGAGCAGTCCGACGCCAAGAGGCTCTTCAAGGAGCAATTTGA 235

Db 3646 TBRNMGEMATVMBRNTTACANTSNAMNCASNMRGMATYDARSARCHANDVMTCT 3705

CC The present invention describes a method for examining ischemic
CC conditions, comprising measuring the expression levels of particular
CC genes (I) in a test sample or determining the expression profile of a
CC gene group in the sample comprising genes selected from (I). The method
CC is useful for examining the ischemic condition (e.g. compressive
CC ischaemia, occlusive ischaemia or vasospastic ischaemia), by measuring
CC expression levels of particular genes (AB199202 to AB199912, encoding
CC the protein sequences in AB199202 to AB199912) or by determining the
CC expression profile of a gene group comprising these genes. The
CC expression levels or expression profiles produced by these genes are
CC used as an indicator when screening for ischemic condition-improving
CC drugs or therapeutics for ischemic diseases. AB199913 and AB199914
CC represent PCR primers for a mouse ischemic condition related sequence,
CC which are used in the exemplification of the present invention.

XX Sequence 1595 BP; 421 A; 350 C; 464 G; 360 T; 0 other;

Query Match 1.6%; Score 42.2; DB 24; Length 1595;
Best Local Similarity 59.7%; Pred. No. 0.057;
Matches 71; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

OY 275 AACCTGTAGAGTCTATGGGTGGGATCGCCCTTAATGTCTCCAAAGATGACCTGG 334
DB 660 ACCCTGGAAGAAATCTTTGTGGAGGTCTAAACCTGAAGCCACAGAGAAAGATCA 719
OY 335 AGGAAGACTTCAGCAAGTTGGGAAATCGAGGATTTAGTTCTCAGAGACGCAAG 393
DB 720 GAGATATCTTCGGCCAGTTGGGAGATTGAGGCCATTGACCTCCAAATGATCCCAAG 778

RESULT 11

AAH14157
ID .AAH14157 standard; cDNA: 1893 BP.

XX AAH14157;

XX 26-JUN-2001 (first entry)

XX Human cDNA sequence SEQ ID NO:11382.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

XX Homo sapiens.

XX EP1074617-A2.

XX 07-FEB-2001.

XX 28-JUL-2000; 2000EP-0116126.

XX 29-JUL-1999; 99JP-0248036.

XX 27-AUG-1999; 99JP-0300253.

XX 11-JAN-2000; 2000JP-0118776.

XX 02-MAY-2000; 2000JP-0183767.

XX 09-JUN-2000; 2000JP-0241899.

XX (HELI-) HELIX RBS INST.

XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI: 2001-318749/34.

XX Primer sets for synthesizing polynucleotides, particularly the 5602

XX full-length cDNAs defined in the specification, and for the detection

XX and/or diagnosis of the abnormality of the proteins encoded by the

XX full-length cDNAs -
XX Claim 8; SEQ ID 11382; 2537bp + CD ROM; English.
XX The present invention describes primer sets for synthesizing 5602
XX full-length cDNAs defined in the specification, where a primer set
XX comprises: (a) an oligo-dT primer and an oligonucleotide complementary

CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.

XX Sequence 1893 BP; 633 A; 305 C; 547 G; 408 T; 0 other;

Query Match 1.6%; Score 42.2; DB 22; Length 1893;
Best Local Similarity 59.7%; Pred. No. 0.063;
Matches 71; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

OY 151 TCTTCACGAGCGTTTGCCTTATATACACAGACATGTCGAGAGACAGTCGCCACCA 210
DB 1355 TCTTCACGAGCGTTTGTGTGTAGACTTCAACAGTCGAGAGATGCCAAGCTGCCAAG 1414

OY 211 GAGGCTCTCAAGGACGCAATTTGAAATGAGTCAATTAAGATGATGACGACGAC 269
DB 1415 GAGGCCATGGAAGACGCGTGAATGATGATAATTAAGTTACCTTGCGGCGCAAC 1473

RESULT 12

ABK63660
ID ABK63660 standard; cDNA: 2142 BP.

XX ABK63660;

XX 18-JUN-2002 (first entry)

XX Rat sequence differentially expressed in response to a hepatotoxin #1567.

XX Rat; ss; hepatotoxin; expressed sequence tag; EST; drug screening;

XX differential expression; centrilobular necrosis; steatosis.

XX Rattus norvegicus.

XX WO200210453-A2.

XX 07-FEB-2002.

XX 30-JUL-2001; 2001WO-US23872.

XX 31-JUL-2000; 2000US-222040P.

XX 02-NOV-2000; 2000US-244880P.

XX 11-MAY-2001; 2001US-290029P.

XX 15-MAY-2001; 2001US-290645P.

XX 22-MAY-2001; 2001US-292336P.

XX 06-JUN-2001; 2001US-295798P.

XX 13-JUN-2001; 2001US-297457P.

XX 19-JUN-2001; 2001US-298884P.

XX 09-JUL-2001; 2001US-303459P.

XX (GENE-) GENE LOGIC INC.

XX Mendrick D, Porter MW, Johnson KR, Castle AL, Elashoff MR;
XX WPI: 2002-241625/29.
XX Predicting toxic effects of compounds or the progression of these toxic

PF 18-MAY-2001; 2001WO-JP04192

human; open reading frame; ORF; detection; cytosolic; hepatotropic; vulnary; antipsoriatic; antiParkinsonian; neurotropic; neuroprotective; anticonvulsant; osteopathic; antiarthritis; immunosuppressant; cardiac; immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive; dermatological; immunosuppressive; antinflammatory; antiviral; antibacterial; antifungal; antineumatic; antihypert; antianemic; gene therapy; cancer; proliferative disorder; hypertension; neurodegenerative disorder; osteoarthritis; graft vs host disease; cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS; cholesterol ester storage; systemic lupus erythematosus; infection; severe combined immunodeficiency; malaria; autoimmune disorder; asthma; allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound; bone damage; cartilage damage; antinflammatory disease; coagulation;

KW thrombosis; contraceptive; ss.
 XX Homo sapiens.
 OS
 XX MO200058473-A2.
 XX
 XX 05-OCT-2000.
 XX
 XX 31-MAR-2000; 2000WO-US08621.
 XX
 XX 31-MAR-1999; 99US-0127607.
 PR 02-APR-1999; 99US-0127636.
 PR 05-APR-1999; 99US-0127728.
 PR 30-MAR-2000; 2000US-0540763.
 XX
 XX (CURA-) CURAGEN CORP.
 PA
 XX Shimkels RA, Leach M;
 PI
 XX WPI; 2000-602362/57.
 DR P-PSDB; AAB42491.
 XX
 XX Novel nucleic acids and peptides derived from open reading frame X,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease -
 XX
 XX Claim 5; Page 3693-3700; 5507pp; English.
 PS
 XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
 CC sequences have activities such as: cytostatic; hepatotropic; vulnery;
 CC antiproliferative; antiparagonian; nocrotic; neuroprotective;
 CC osteopathic; anticonvulsant; antiarthritic; immunosuppressive;
 CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
 CC antidiabetic; hypotensive; dermatological; immunosuppressive;
 CC antitumor; antitubercular; antiviral; antifungal; antirheumatic;
 CC antihypertoid; and antianemic. The sequences can be used for determining
 CC the presence of or predisposition to, or preventing or treating
 CC pathological conditions associated with an ORFX-associated disorder. The
 CC nucleic acids can be used to express ORFX proteins in gene therapy
 CC vectors. The proteins and nucleic acids may be used to treat cancers,
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,
 CC hyperension, hypothyroidism, cholesterol ester storage, systemic lupus
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
 CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
 CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance
 CC coagulation; to inhibit thrombosis; and as a contraceptive.
 XX
 XX Sequence 11680 BP; 3290 A; 3113 C; 2858 G; 2417 T; 2 other;
 SQ
 Query Match 1.5%; Score 41; DB 21; Length 11680;
 Best Local Similarity 48.9%; Pred. No. 0.49;
 Matches 110; Conservative 0; Mismatches 115; Indels 0; Gaps 0;

RESULT 15
 ABL33124/C
 ID ABL33124 standard; DNA; 6104 BP.
 XX
 XX ABL33124;
 AC
 XX 26-MAR-2002 (first entry)
 XX
 XX Human immune system associated gene SEQ ID NO: 1097.
 DE
 XX Human; immune system disease; cytosine methylation; antileptomatic;
 KW antileptomatic; antianemic; cytosine methylation; nocrotic;
 KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
 KW antirheumatic; antiarthritic; antidiabetic; antiproliferative;
 KW antitubercular; cancer; eye disease; arteriosclerosis; anaemia;
 KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
 KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
 KW gene; ds.
 XX
 XX Homo sapiens.
 OS
 XX
 XX WO200200928-A2.
 PN
 XX 03-JAN-2002.
 PD
 XX 02-JUL-2001; 2001WO-EP07537.
 PF
 XX 30-JUN-2000; 2000DE-1032529.
 PR 01-SEP-2000; 2000DE-1043826.
 PR
 XX (EPIC-) EPICENOMICS AG.
 PA
 XX Olek A, Piepenbrock C, Berlin K;
 PI
 XX WPI; 2002-130909/17.
 DR
 XX Nucleic acid comprising fragment of chemically modified gene, useful
 PT for diagnosis and treatment of diseases associated with abnormal
 PT cytosine methylation -
 XX
 XX Claim 1; SEQ ID NO 1097; 32bp + Sequence Listing; German.
 PS
 XX The present invention provides a number of human immune system associated
 CC genes which are modified by the methylation of cytosines. The sequences
 CC can be used in the diagnosis and treatment of immune system disorders,
 CC including eye diseases such as retinopathy, neovascular glaucoma and
 CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
 CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
 CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
 CC diseases. The present sequence is a gene of the invention.
 XX
 XX Sequence 6104 BP; 1398 A; 129 C; 1826 G; 2751 T; 0 other;
 SQ
 Query Match 1.5%; Score 39.6; DB 24; Length 6104;
 Best Local Similarity 51.7%; Pred. No. 0.89;
 Matches 90; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

Search completed: June 19, 2003, 01:26:17
 Job time : 612 secs

This Page Blank (uspto)